



PCT

特許協力条約に基づいて公開された国際出願

<p>(51) 国際特許分類6</p> <p style="text-align: center;">G06F 17/30</p>	A1	<p>(11) 国際公開番号</p> <p style="text-align: right;">WO96/29659</p> <p>(43) 国際公開日</p> <p style="text-align: right;">1996年9月26日(26.09.96)</p>												
<p>(21) 国際出願番号 PCT/JP96/00709</p> <p>(22) 国際出願日 1996年3月18日(18.03.96)</p> <p>(30) 優先権データ</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 40%;">特願平7/59196</td> <td style="width: 40%;">1995年3月17日(17.03.95)</td> <td style="width: 20%; text-align: center;">JP</td> </tr> <tr> <td>特願平7/59200</td> <td>1995年3月17日(17.03.95)</td> <td style="text-align: center;">JP</td> </tr> <tr> <td>特願平7/106173</td> <td>1995年4月28日(28.04.95)</td> <td style="text-align: center;">JP</td> </tr> <tr> <td>特願平7/106181</td> <td>1995年4月28日(28.04.95)</td> <td style="text-align: center;">JP</td> </tr> </table> <p>(71) 出願人 (米国を除くすべての指定国について)</p> <p>呉羽化学工業株式会社</p> <p>(KUREHA KAGAKU KOGYO KABUSHIKI KAISHA)(JP/JP)</p> <p>〒103 東京都中央区日本橋堀留町一丁目9番11号 Tokyo, (JP)</p> <p>(72) 発明者: および</p> <p>(75) 発明者/出願人 (米国についてのみ)</p> <p>朝永 博(TOMONAGA, Atsushi)(JP/JP)</p> <p>〒180 東京都武蔵野市境南町三丁目17-13 Tokyo, (JP)</p> <p>田村文男(TAMURA, Fumio)(JP/JP)</p> <p>〒300-12 茨城県稲敷郡笠崎町高見原三丁目3-51 Ibaraki, (JP)</p>		特願平7/59196	1995年3月17日(17.03.95)	JP	特願平7/59200	1995年3月17日(17.03.95)	JP	特願平7/106173	1995年4月28日(28.04.95)	JP	特願平7/106181	1995年4月28日(28.04.95)	JP	<p>(74) 代理人</p> <p>弁理士 長谷川芳樹, 外(HASEGAWA, Yoshiki et al.)</p> <p>〒104 東京都中央区京橋二丁目13番10号</p> <p>京橋ナショナルビル6F 創英国際特許事務所 Tokyo, (JP)</p> <p>(81) 指定国</p> <p>AU, CA, CN, KR, RU, US, 欧州特許(AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p style="text-align: center;">添付公開書類 国際調査報告書</p>
特願平7/59196	1995年3月17日(17.03.95)	JP												
特願平7/59200	1995年3月17日(17.03.95)	JP												
特願平7/106173	1995年4月28日(28.04.95)	JP												
特願平7/106181	1995年4月28日(28.04.95)	JP												
<p>(54) Title : BIOCHEMICAL INFORMATION PROCESSOR, BIOCHEMICAL INFORMATION PROCESSING METHOD, AND BIOCHEMICAL INFORMATION RECORDING MEDIUM</p>														
<p>(54) 発明の名称 生化学情報処理装置及び生化学情報処理方法</p>														
<p>(57) Abstract</p> <p>A biochemical information processor provided with means for storing biochemical information, means for inputting data, means for detecting chemical reaction formulae related to a compound on the basis of the data, and means for displaying a chart of chemical reaction formulae. The storage means is provided with a compound information file, an enzyme information file, and an interrelationship information file, the last file storing a table indicating the interrelationship among the number of a compound serving as a key, the number of an enzyme against which the compound serves as a substrate, and the number of an enzyme against which the compound serves as a product. The reaction formula detection means is provided with a first processing section for preparing normalized data of the compound from the above data and retrieving the compound information file on the basis of normalized data to thereby reading out the compound number, a second processing section for reading out the number of the enzyme against which the compound serves as a substrate or a product from the interrelationship information file on the basis of the compound number read out in the first section, a third processing section for reading out the number of each of other compound(s) constituting a reaction system with the enzyme and additional information about the enzyme from the enzyme information file, and a fourth processing section for displaying the chart of reaction formulae of the compound on the basis of the compound and enzyme numbers.</p>														
<div style="display: flex; justify-content: space-between;"> <div style="width: 60%;"> <p>4 ... Biochemical information processing program</p> <p>5 ... 1st storage unit</p> <p>6 ... 2nd storage unit</p> <p>7 ... Input unit</p> <p>8 ... Image memory</p> <p>9 ... Working memory</p> <p>10 ... Main program</p> <p>11 ... Three-dimensional display program</p> <p>12 ... Reaction formula detecting program</p> <p>13a ... 1st processing routine</p> <p>13b ... 2nd processing routine</p> <p>13c ... 3rd processing routine</p> <p>13d ... 4th processing routine</p> <p>14 ... Receptor information detecting program</p> <p>15a ... 5th processing routine</p> <p>15b ... 6th processing routine</p> <p>15c ... 7th processing routine</p> <p>15d ... 8th processing routine</p> <p>16 ... Reaction path detecting program</p> <p>17a ... 9th processing routine</p> <p>17b ... 10th processing routine</p> <p>17c ... 11th processing routine</p> <p>17d ... 12th processing routine</p> <p>17e ... 13th processing routine</p> <p>17f ... 14th processing routine</p> <p>18 ... Compound information file</p> <p>19 ... Enzyme information file</p> <p>20 ... Interrelationship information file</p> <p>21 ... Partial interrelationship data file</p> <p>22 ... Name table information file</p> <p>23 ... Receptor information file</p> <p>24 ... Display</p> <p>25 ... Mouse</p> <p>26 ... Keyboard</p> <p>27 ... Printer</p> </div> <div style="width: 35%;"> <p style="text-align: center;">BIOCHEMICAL INFORMATION PROCESSOR</p> </div> </div>														

(57) 要約

生化学情報が記憶された記憶手段と、データの入力を受け付ける入力手段と、データに基づいて化合物が関わる化学反応式を検出する反応式検出手段と、化学反応式の反応式図を表示する表示手段とを備える生化学情報処理装置である。前記記憶手段は、化合物情報ファイルと、酵素情報ファイルと、相互関連情報ファイルとを備えており、相互関連情報ファイルには、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化合物が生成物となる酵素の番号との関係を示す一覧表が格納されている。また、前記反応式検出手段は、データからその化合物の正準化データを作成し、それに基づいて化合物情報ファイルを検索して化合物の番号を読み出す第1の処理部と、化合物の番号に基づいて化合物が基質あるいは生成物となる酵素の番号を相互関連情報ファイルから読み出す第2の処理部と、酵素と反応系を構成する他の化合物の番号及び酵素の付加情報を酵素情報ファイルから読み出す第3の処理部と、化合物の番号と酵素の番号とからこの化合物の反応式図を前記表示手段に表示する第4の処理部とを備える。

情報としての用途のみ

PCTに基づいて公開される国際出願をパンフレット第一頁にPCT加盟国を特定するために使用されるコード

AL アルバニア
AM アルメニア
AU オーストラリア
AZ アゼルバイジャン
BA ボスニア・ヘルツェゴビナ
BB ベルバドス
BE ベルギー
BF ブルキナ・ファソ
BG ブルガリア
BJ ベナン
BR ブラジル
BY ベラルーシ
CA カナダ
CF 中央アフリカ共和国
CG コンゴ
CH スイス
CI コート・ジボアール
CM カメルーン
CN 中国
CU キューバ
CZ チェコ共和国

DE ドイツ
DK デンマーク
EE エストニア
ES スペイン
FI フィンランド
FR フランス
GB ガブリウス
GR ギリシア
GN ギニア
GR ギリシャ
HU ハンガリー
IE アイルランド
IL イスラエル
IS アイスラ
IT イタリア
JP 日本
KE ケニア
KG キルギスタン
KP 朝鮮民主主義人民共和国
KR 大韓民国
KZ カザフスタン

LI リヒテンシュタイン
LC セントルシア
LK スリランカ
LR リベリア
LS レソト
LT リトアニア
LU ルクセンブルグ
LV ラトヴィア
MC モナコ
MD モルドヴァ共和国
MG マダガスカル
MK マケドニア共和国
ML マリ
MN モンゴル
MR モリタニア
MW マラウイ
MX メキシコ
NE ニジェール
NL オランダ
NO ノルウェー
NZ ニュージーランド

PL ポーランド
PT ポルトガル
RU ロシア連邦
SE スウェーデン
SG シンガポール
SI スロベニア
SK スロバキア
SN セネガル
SZ スワジランド
TD チャド
TG トーゴ
TH タイ
TJ タジキスタン
TM トルクメニスタン
TN トンガ
TR トルク
TT トリニダード・トバゴ
UA ウクライナ
UG ウガンダ
US アメリカ合衆国
UZ ウズベキスタン
VN ヴェトナム

明 細 書

生化学情報処理装置及び生化学 報処理方法

技術分野

本発明は、生化学分野の情報処理を行う処理装置及び処理方法に関し、特に、生体関連化合物の反応経路の検索、及びこの反応経路の連続的表示を行うと共に、生体関連物質に関する情報を得ることのできる処理装置及び処理方法に関する。

また、本発明は、生化学情報を記録したフレキシブルディスク、磁気テープ等の情報記録媒体（コンピュータプログラム製品）に関し、特に、生体関連化合物の反応経路の検索を行うための情報、この反応経路の連続的表示を行うための情報、生体関連物質に関する情報等が記録された情報記録媒体に関する。

背景技術

従来より、化合物情報を収載した化合物データベースシステムやプログラム、及び化合物の反応情報を収載した反応データベースシステムやプログラムが開発されている。化合物データベースシステムやプログラムには、既存の化合物の物性や作用などの化合物情報が収載されており、化合物の構造をキーとして化合物情報にアクセスするものである。また、反応データベースシステムには、既存の化合物の反応情報が収載されており、化合物の構造をキーとして反応情報にアクセスするものである。

このような化合物データベースとしては、例えば、米国MDL社の化合物管理システム“MACCS”がある。また、反応データベースシステムとしては、例えば、米国MDL社の総合化学情報管理システム“ISIS”や、反応情報管理システム“REACCS”がある。

しかしながら、従来の化合物／反応データベースシステムには、化合物と酵素

との関係及び生体関連物質に関する情報が、統合的に収載されたものが存在しない。このため、化合物の構造をキーとしたのでは、酵素についての情報、あるいは、酵素・基質・生成物に関連した生化学情報を効率的に得ることができなかった。また、従来の化合物／反応データベースシステムには、複数の化合物の反応経路が統合的に構築されたものが存在しない。このため、効率的に複数の化合物がかかわる反応経路を検索することができなかった。

さらに、従来の化合物／反応データベースシステムには、生体機能の調節や生体内での情報伝達のために存在する受容体に関する情報、及び生体関連物質（作動物質・拮抗物質）に関する情報が、総合的に収載されたものが存在しない。このため、効率的に受容体・作動物質・拮抗物質に関連した生化学情報を得ることができなかった。

本発明は、このような問題を解決し、化合物構造をキーとした場合でも、酵素についての情報、あるいは、酵素・基質・生成物に関連した生化学情報を効率的に得ることができ、また複数の化合物が関わる反応経路を効率的に検索することができ、さらに受容体・作動物質・拮抗物質に関連した生化学情報を効率的に得ることができる生化学情報処理装置、生化学情報処理方法及び情報記録媒体（コンピュータプログラム製品）を提供することを目的とする。

発明の開示

先ず、本発明の生化学情報処理装置について説明する。

本発明の生化学情報処理装置は、

化合物と酵素とについての生化学情報が記憶された記憶手段と、

前記生化学情報を表す画像データまたは前記生化学情報を表す記号データの入力を受け付ける入力手段と、

基質および／または生成物である化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物が関わる化学反応式を検

出する反応式検出手段と、

少なくとも化学反応式の反応式図を表示する表示手段とを備える生化学情報処理装置であって；

前記記憶手段は、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及び当該化合物についての付加情報が格納された化合物情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物である化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納された酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化合物が生成物となる酵素の番号との関係を示す一覧表が格納された相互関連情報ファイルとを備え；

前記反応式検出手段は、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらに正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第1の処理部と、

前記第1の処理部で読み出された化合物の番号に基づいて、化合物が基質あるいは生成物となる酵素の番号を前記相互関連情報ファイルから読み出す第2の処理部と、

前記第2の処理部で読み出された番号の酵素と前記化合物と共に反応系を構成する他の化合物の番号、及び前記酵素についての付加情報を前記酵素情報ファイルから読み出す第3の処理部と、

前記第1の処理部で読み出された化合物の番号と、前記第2の処理部で読み出された酵素の番号と、前記第3の処理部で読み出された他の化合物の番号とから、

前記入力手段で受入れられた化合物の反応式図を前記表示手段に表示し、さらに、前記第3の処理部で読み出された酵素についての付加情報を前記表示手段に表示する第4の処理部とを備える。

上記本発明の生化学情報処理装置によれば、入力手段で受け付けられた化合物についてのデータが第1の処理部に与えられると、このデータから正準化データが作成される。そして、作成された正準化データに基づいて化合物情報ファイルが検索され、正準化データが化合物情報ファイルに存在する場合には、正準化データに対応する化合物の番号が読み出される。第1の処理部で読み出された化合物の番号は第2の処理部に与えられ、第2の処理部では、この化合物が基質あるいは生成物となる酵素の番号が相互関連情報ファイルから読み出される。

第2の処理部で読み出された酵素の番号は第3の処理部に与えられ、第3の処理部では、この酵素と前記化合物と共に反応系を構成する他の化合物の番号、及び前記酵素についての付加情報が酵素情報ファイルから読み出される。そして、第1の処理部で読み出された化合物の番号と、第2の処理部で読み出された酵素の番号と、前記第3の処理部で読み出された他の化合物の番号とが第4の処理部に与えられ、第4の処理部では、前記入力手段で受け付けられた化合物の反応式図を表示手段に表示させる。同様に、第3の処理部で読み出された酵素についての付加情報も表示手段に表示させる。

また、本発明の生化学情報処理装置は、化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物が作動物質および／または拮抗物質となる受容体についての付加情報を検出する受容体情報検出手段を更に備えていてもよく、この場合；

前記記憶手段が、受容体についての生化学情報を更に記憶しており、かつ

受容体の番号と当該受容体の作動物質および／または拮抗物質である化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された受容体情報ファイルを更に備えており；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表を格納しており；

前記受容体情報検出手段は、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第5の処理部と、

前記第5の処理部で読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す第6の処理部と、

前記第6の処理部で読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す第7の処理部と、

前記第7の処理部で読み出された受容体についての少なくとも付加情報を前記表示手段に表示する第8の処理部とを備える。

この場合、本発明の生化学情報処理装置においては、入力手段で受け付けられた化合物についてのデータが第5の処理部に与えられると、このデータから正準化データが作成される。そして、作成された正準化データに基づいて化合物情報ファイルが検索され、正準化データが化合物情報ファイルに存在する場合には、正準化データに対応する化合物の番号が読み出される。第5の処理部で読み出された化合物の番号は第6の処理部に与えられ、第6の処理部では、この化合物が作動物質あるいは拮抗物質となる受容体の番号が相互関連情報ファイルから読み出される。第6の処理部で読み出された受容体の番号は第7の処理部に与えられ、第7の処理部では、この受容体についての少なくとも付加情報が受容体情報ファ

イルから読み出される。そして、第7の処理部で読み出された受容体についての少なくとも付加情報が第8の処理部に与えられ、第8の処理部では、この受容体についての少なくとも付加情報を表示手段に表示させる。

また、本発明の生化学情報処理装置は、反応経路を構成する複数の化合物の中から選択される所定の化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記複数の化合物の反応経路を検出する反応経路検出手段を更に備えていてもよく、この場合；

前記反応経路検出手段は、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第9の処理部と、

前記第9の処理部で読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す第10の処理部と、

前記第10の処理部で読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報ファイルから読み出す第11の処理部と、

前記第10の処理部による処理と前記第11の処理部による処理とを繰り返し、所定反応経路内の化合物及び酵素を検索する第12の処理部と、

前記第10の処理部で読み出された酵素の番号と、前記第11の処理部で読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前記表示手段に表示する第13の処理部とを備える。

この場合、本発明の生化学情報処理装置においては、入力手段で受け付けられた化合物についてのデータが第9の処理部に与えられると、このデータから正準

化データが作成される。そして、作成された正準化データに基づいて化合物情報ファイルが検索され、正準化データが化合物情報ファイルに存在する場合には、正準化データに対応する化合物の番号が読み出される。第9の処理部で読み出された化合物の番号は第10の処理部に与えられ、第10の処理部では、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とが、相互関連情報ファイルから読み出される。

第10の処理部で読み出された酵素の番号は第11の処理部に与えられ、第11の処理部では、この酵素について、基質である化合物の番号とこの酵素による生成物である化合物の番号とが、酵素情報ファイルから読み出される。第10の処理部と第11の処理部の各処理は、第12の処理部によって繰り返される。

そして、第10の処理部で読み出された酵素の番号と、第11の処理部で読み出された化合物の番号とは第13の処理部に与えられ、第13の処理部では、これらの化合物の反応式図を所定の反応経路に沿って表示手段に表示させる。

更に、本発明の生化学情報処理装置は以下のものであってもよい。すなわち、化合物と酵素とについての生化学情報が記憶された記憶手段と、

前記生化学情報を表す画像データまたは前記生化学情報を表す記号データの入力を受け付ける入力手段と、

反応経路を構成する複数の化合物の中から選択される所定の化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記複数の化合物の反応経路を検出する反応経路検出手段と、

少なくとも化学反応式の反応式図を表示する表示手段とを備える生化学情報処理装置であって；

前記記憶手段は、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及び当該化合物についての付加情報が格納された化合物情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物であ

る化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納された酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化合物が生成物となる酵素の番号との関係を示す一覧表が格納された相互関連情報ファイルとを備え；

前記反応経路検出手段は、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第9の処理部と、

前記第9の処理部で読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す第10の処理部と、

前記第10の処理部で読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報ファイルから読み出す第11の処理部と、

前記第10の処理部による処理と前記第11の処理部による処理とを繰り返し、所定反応経路内の化合物及び酵素を検索する第12の処理部と、

前記第10の処理部で読み出された酵素の番号と、前記第11の処理部で読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前記表示手段に表示する第13の処理部とを備える。

この場合、本発明の生化学情報処理装置は、化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物が作動物質および／または拮抗物質となる受容体についての付加情報を検出する受容体情報検出手段を更に備えていてもよく、この場合；

前記記憶手段が、受容体についての生化学情報を更に記憶しており、かつ受容体の番号と当該受容体の作動物質および／または拮抗物質である化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された受容体情報ファイルを更に備えており；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表を格納しており；

前記受容体情報検出手段は、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第5の処理部と、

前記第5の処理部で読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す第6の処理部と、

前記第6の処理部で読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す第7の処理部と、

前記第7の処理部で読み出された受容体についての少なくとも付加情報を前記表示手段に表示する第8の処理部とを備える。

更に、本発明の生化学情報処理装置においては、

前記入力手段が、化合物を構成する各原子についての固有データ及び原子間の結合対データの入力を受け付けるものであり；

前記生化学情報処理装置が、前記入力手段で受け付けられた各データに基づいて、前記化合物の化学構造を一意的に特定できる正準化データを作成する以下の

正準化データ作成手段を更に備えていることが好ましい。すなわち、前記正準化データ作成手段は、

前記入力手段で受け付けられた各データに基づいて、各原子を等価原子ごとに別のクラスに分類して、クラスごとに異なるクラス番号を各原子に与える構成原子分類処理部と、

前記構成原子分類処理部で各原子に与えられたクラス番号に基づいて、前記化合物の構造と一意的に対応した正準化番号を各原子に与える正準化番号付与処理部と、

前記正準化番号付与処理部で各原子に与えられた正準化番号に基づいて、前記正準化データを作成する正準化データ作成処理部とを備える。

このような構成を有する本発明にかかる正準化データ作成手段によれば、入力手段で受け付けられた各原子についての固有データ及び原子間の結合対データは正準化データ作成手段に与えられる。そして、正準化データ作成手段では、これらのデータに基づいて正準化データが作成される。

即ち、正準化データ作成手段では、まず、構成原子分類処理部の処理を実行し、各原子についての固有データ及び原子間の結合対データに基づいて、各原子を等価原子ごとに別のクラスに分類する。そして、クラスごとに異なるクラス番号を各原子に与える。次に、正準化番号付与処理部の処理を実行し、各原子に与えられたクラス番号及び原子間の結合対データに基づいて、化合物の構造と一意的に対応した正準化番号を各原子に与える。さらに、正準化データ作成処理部の処理を実行し、各原子に与えられた正準化番号及び各原子についての固有データに基づいて正準化データを作成する。

ここで、前記構成原子分類処理部は、

各原子に3種類の属性 (a_i , b_{ij} , d_{ij}) を与え、これらの属性が一つでも異なる原子は非等価であると判定できることを利用して、各原子を等価原子毎に異なるクラス番号を付与しており、

前記3種類の属性 (a_i , b_{ij} , d_{ij}) の中で、 a_i は入力番号 i の原子の種類番号であり、 b_{ij} は入力番号 i の原子に隣接する結合のうち、その種類番号が j である結合の数であり、 d_{ij} は入力番号 i の原子から最短経路により j 個の結合を経て巡れる道筋の数であり；

前記正準化番号付与処理部は、

正準化番号を1から昇順に各原子に与える過程において、前記クラス番号の優先順位が最高である原子に正準化番号1を与え、以降正準化番号 n までが付与されている時、既に正準化番号が与えられている原子でかつ未だ正準化番号が与えられていない原子が結合している原子の中で正準化番号が最小である原子を選び、その原子に結合している未だ正準化番号が与えられていない原子の中で前記クラス番号の優先順位が最高である原子に正準化番号 $n+1$ を与えており；

前記正準化データ作成処理部は、

各原子に3種類の属性 (P_i , T_i , S_i) を与えて、これらの属性を一列に並べることによって前記正準化データを作成しており、

前記3種類の属性 (P_i , T_i , S_i) の中で、 P_i は正準化番号 i の原子に結合し且つ正準化番号が最小の原子の正準化番号であり、 T_i は正準化番号 i の原子と正準化番号 P_i の原子との結合の種類記号であり、 S_i は正準化番号 i の原子の種類記号であることが好ましい。

次に、本発明の生化学情報処理方法について説明する。

本発明の生化学情報処理方法は、

化合物と酵素とについての生化学情報が記憶された記憶手段と、

前記生化学情報を表す画像データまたは前記生化学情報を表す記号データの入力を受け付ける入力手段と、

少なくとも化学反応式の反応式図を表示する表示手段とを備える情報処理装置を用いた生化学情報処理方法であって；

前記記憶手段は、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及び当該化合物についての付加情報が格納された化合物情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物である化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納された酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化合物が生成物となる酵素の番号との関係を示す一覧表が格納された相互関連情報ファイルとを備え；

前記生化学情報処理方法は、

基質および／または生成物である化合物についてのデータが前記入力手段で受け付けられた場合に、このデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第1のステップと、

前記第1のステップで読み出された化合物の番号に基づいて、この化合物が基質あるいは生成物となる酵素の番号を前記相互関連情報ファイルから読み出す第2のステップと、

前記第2のステップで読み出された番号の酵素と前記化合物と共に反応系を構成する他の化合物の番号、及び前記酵素についての付加情報を、前記酵素情報ファイルから読み出す第3のステップと、

前記第1のステップで読み出された化合物の番号と、前記第2のステップで読み出された酵素の番号と、前記第3のステップで読み出された他の化合物の番号とから、前記入力手段で受け付けられた化合物の反応式図を前記表示手段に表示し、さらに、前記第3のステップで読み出された酵素についての付加情報を前記表示手段に表示する第4のステップとを備える。

上記本発明の生化学情報処理方法によれば、第1のステップから第4のステッ

ブまでの処理によって、反応式の検出を行うことができる。反応式の検出は、まず、第1のステップの処理が行われ、入力手段で受け付けられた化合物についてのデータから正準化データが作成される。そして、作成された正準化データに基づいて化合物情報ファイルが検索され、正準化データが化合物情報ファイルに存在する場合には、正準化データに対応する化合物の番号が読み出される。次に、第2のステップの処理が行われ、第1のステップで読み出された化合物の番号に基づいて、この化合物が基質あるいは生成物となる酵素の番号が、相互関連情報ファイルから読み出される。

さらに、第3のステップの処理が行われ、第2のステップで読み出された番号の酵素と前記化合物と共に反応系を構成する他の化合物の番号、及び前記酵素についての付加情報が酵素情報ファイルから読み出される。そして、第4のステップの処理が行われ、第1のステップで読み出された化合物の番号と、第2のステップで読み出された酵素の番号と、前記第3のステップで読み出された他の化合物の番号とから、前記入力手段で受け付けられた化合物の反応式図を表示手段に表示させる。同様に、第3のステップで読み出された酵素についての付加情報も表示手段に表示させる。

また、本発明の生化学情報処理方法においては、

前記記憶手段が、受容体についての生化学情報を更に記憶しており、かつ

受容体の番号と当該受容体の作動物質および／または拮抗物質である化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された受容体情報ファイルを更に備えていてもよく、この場合；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表を格納しており；

前記生化学情報処理方法が、

化合物についてのデータが前記入力手段で受け付けられた場合に、このデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第5のステップと、

前記第5のステップで読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す第6のステップと、

前記第6のステップで読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す第7のステップと、

前記第7のステップで読み出された受容体についての少なくとも付加情報を前記表示手段に表示する第8のステップとを更に備える。

この場合、本発明の生化学情報処理方法においては、第5のステップから第8のステップまでの処理によって、受容体情報の検出を行うことができる。受容体情報の検出は、まず、第5のステップの処理が行われ、入力手段で受け付けられた化合物についてのデータから正準化データが作成される。そして、作成された正準化データに基づいて化合物情報ファイルが検索され、正準化データが化合物情報ファイルに存在する場合には、正準化データに対応する化合物の番号が読み出される。次に、第6のステップの処理が行われ、第5のステップで読み出された化合物の番号に基づいて、この化合物が作動物質あるいは拮抗物質となる受容体の番号が、相互関連情報ファイルから読み出される。さらに、第7のステップの処理が行われ、第6のステップで読み出された番号の受容体についての少なくとも付加情報が受容体情報ファイルから読み出される。そして、第8のステップの処理が行われ、第7のステップで読み出された受容体についての少なくとも付加情報を表示手段に表示させる。

また、本発明の生化学情報処理方法は、

反応経路を構成する複数の化合物の中から選択される所定の化合物についてのデータが前記入力手段で受け付けられた場合に、このデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第9のステップと、

前記第9のステップで読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す第10のステップと、

前記第10のステップで読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報ファイルから読み出す第11のステップと、

前記第10のステップと前記第11のステップとを繰り返し、所定反応経路内の化合物及び酵素を検索する第12のステップと、

前記第10のステップで読み出された酵素の番号と、前記第11のステップで読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前記表示手段に表示する第13のステップとを更に備えていてもよい。

この場合、本発明の生化学情報処理方法においては、第9のステップから第12のステップまでの処理によって反応経路の検出を行うことができる。反応経路の検出は、まず、第9のステップの処理が行われ、入力手段で受け付けられた所定の化合物についてのデータから正準化データが作成される。そして、作成された正準化データに基づいて化合物情報ファイルが検索され、正準化データが化合物情報ファイルに存在する場合には、正準化データに対応する化合物の番号が読み出される。次に、第10のステップの処理が行われ、第9のステップで読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とが、相互関連情報ファイルから読み出される。

さらに、第 11 のステップの処理が行われ、第 10 のステップで読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号とが、酵素情報ファイルから読み出される。第 10 のステップと第 11 のステップとの各処理は、第 12 のステップによって繰り返される。

そして、第 13 のステップの処理が行われ、第 10 のステップで読み出された酵素の番号と、第 11 のステップで読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って表示手段に表示させる。

更に、本発明の生化学情報処理方法は以下のものであってもよい。すなわち、化合物と酵素とについての生化学情報が記憶された記憶手段と、

前記生化学情報を表す画像データまたは前記生化学情報を表す記号データの入力を受け付ける入力手段と、

少なくとも化学反応式の反応式図を表示する表示手段とを備える情報処理装置を用いた生化学情報処理方法であって；

前記記憶手段は、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及び当該化合物についての付加情報が格納された化合物情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物である化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納された酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化合物が生成物となる酵素の番号との関係を示す一覧表が格納された相互関連情報ファイルとを備え；

前記生化学情報処理方法は、

反応経路を構成する複数の化合物の中から選択される所定の化合物についてのデータが前記入力手段で受け付けられた場合に、このデータから当該化合物の化

学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第9のステップと、

前記第9のステップで読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す第10のステップと、

前記第10のステップで読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報ファイルから読み出す第11のステップと、

前記第10のステップと前記第11のステップとを繰り返し、所定反応経路内の化合物及び酵素を検索する第12のステップと、

前記第10のステップで読み出された酵素の番号と、前記第11のステップで読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前記表示手段に表示する第13のステップとを備える。

この場合、本発明の生化学情報処理方法においては、

前記記憶手段が、受容体についての生化学情報を更に記憶しており、かつ

受容体の番号と当該受容体の作動物質および／または拮抗物質である化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された受容体情報ファイルを更に備えていてもよく、この場合；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表を格納しており；

前記生化学情報処理方法が、

化合物についてのデータが前記入力手段で受け付けられた場合に、このデータ

から当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第5のステップと、

前記第5のステップで読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す第6のステップと、

前記第6のステップで読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す第7のステップと、

前記第7のステップで読み出された受容体についての少なくとも付加情報を前記表示手段に表示する第8のステップとを更に備える。

更に、本発明の生化学情報処理方法においては、

前記入力手段が、化合物を構成する各原子についての固有データ及び原子間の結合対データの入力を受け付け；

前記生化学情報処理方法が、

前記入力手段で受け付けられた各データに基づいて、各原子を等価原子ごとに別のクラスに分類して、クラスごとに異なるクラス番号を各原子に与える構成原子分類ステップと、

前記構成原子分類ステップで各原子に与えられたクラス番号に基づいて、前記化合物の構造と一意的に対応した正準化番号を各原子に与える正準化番号付与ステップと、

前記正準化番号付与ステップで各原子に与えられた正準化番号に基づいて、前記化合物の化学構造を一意的に特定できる正準化データを作成する正準化データ作成ステップとを更に備えていることが好ましい。

このような構成を有する本発明にかかる正準化データを作成するための諸ステップによれば、入力手段で受け付けられた各原子についての固有データ及び原子

間の結合対データに基づいて正準化データが作成される。

即ち、まず、構成原子分類ステップにおいて、各原子についての固有データ及び原子間の結合対データに基づいて、各原子を等価原子ごとに別のクラスに分類する。そして、クラスごとに異なるクラス番号を各原子に与える。次に、正準化番号付与ステップにおいて、各原子に与えられたクラス番号及び原子間の結合対データに基づいて、化合物の構造と一意的に対応した正準化番号を各原子に与える。さらに、正準化データ作成ステップにおいて、各原子に与えられた正準化番号及び各原子についての固有データに基づいて正準化データを作成する。

ここで、前記構成原子分類ステップは、

各原子に3種類の属性 (a_i , b_{ij} , d_{ij}) を与え、これらの属性が一つでも異なる原子は非等価であると判定できることを利用して、各原子を等価原子毎に異なるクラス番号を付与しており、

前記3種類の属性 (a_i , b_{ij} , d_{ij}) の中で、 a_i は入力番号 i の原子の種類番号であり、 b_{ij} は入力番号 i の原子に隣接する結合のうち、その種類番号が j である結合の数であり、 d_{ij} は入力番号 i の原子から最短経路により j 個の結合を経て巡れる道筋の数であり；

前記正準化番号付与ステップは、

正準化番号を1から昇順に各原子に与える過程において、前記クラス番号の優先順位が最高である原子に正準化番号1を与え、以降正準化番号 n までが付与されている時、既に正準化番号が与えられている原子でかつ未だ正準化番号が与えられていない原子が結合している原子の中で正準化番号が最小である原子を選び、その原子に結合している未だ正準化番号が与えられていない原子の中で前記クラス番号の優先順位が最高である原子に正準化番号 $n+1$ を与えており；

前記正準化データ作成ステップは、

各原子に3種類の属性 (P_i , T_i , S_i) を与えて、これらの属性を一行に並べることによって前記正準化データを作成しており、

前記3種類の属性 (P_i , T_i , S_i) の中で、 P_i は正準化番号 i の原子に結合し且つ正準化番号が最小の原子の正準化番号であり、 T_i は正準化番号 i の原子と正準化番号 P_i の原子との結合の種類記号であり、 S_i は正準化番号 i の原子の種類記号であることが好ましい。

次に、本発明の生化学情報コンピュータプログラム製品 (生化学情報記録媒体) について説明する。

本発明の生化学情報コンピュータプログラム製品は、

生化学情報を表す画像データまたは生化学情報を表す記号データの入力を受け付ける入力手段と、少なくとも化学反応式の反応式図を表示する表示手段と、コンピュータ使用可能な媒体から情報を読み出す読出手段とを備える情報処理装置と共に用いるための生化学情報コンピュータプログラム製品であって；

ファイルを記録するためのファイル領域とプログラムを記録するためのプログラム領域とを有しており、かつ、前記入力手段から入力されたデータに基づいて少なくとも反応式図を効率的に検索させかつ前記表示手段で表示させるための該媒体中に具現化されたコンピュータ読取可能なファイル及びプログラムを有しているコンピュータ使用可能な媒体を前記コンピュータプログラム製品は備えており；

該コンピュータプログラム製品は、

前記ファイル領域において、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及び当該化合物についての付加情報が格納された、コンピュータ読取可能な化合物情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物である化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納された、コンピュータ読取可能な酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化

化合物が生成物となる酵素の番号との関係を示す一覧表が格納された、コンピュータ読取可能な相互関連情報ファイルとを有しており、

前記プログラム領域において、

基質および／または生成物である化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物に関わる化学反応式を検出する、コンピュータ読取可能な反応式検出プログラムを有しており；

前記反応式検出プログラムは、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらに正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す、コンピュータ読取可能な第 1 の処理ルーチンと、

前記第 1 の処理ルーチンで読み出された化合物の番号に基づいて、化合物が基質あるいは生成物となる酵素の番号を前記相互関連情報ファイルから読み出す、コンピュータ読取可能な第 2 の処理ルーチンと、

前記第 2 の処理ルーチンで読み出された番号の酵素と前記化合物と共に反応系を構成する他の化合物の番号、及び前記酵素についての付加情報を前記酵素情報ファイルから読み出す、コンピュータ読取可能な第 3 の処理ルーチンと、

前記第 1 の処理ルーチンで読み出された化合物の番号と、前記第 2 の処理ルーチンで読み出された酵素の番号と、前記第 3 の処理部で読み出された他の化合物の番号とから、前記入力手段で受け付けられた化合物の反応式図を前記表示手段に表示し、さらに、前記第 3 の処理ルーチンで読み出された酵素についての付加情報を前記表示手段に表示する、コンピュータ読取可能な第 4 の処理ルーチンとを備える。

上記本発明の生化学情報コンピュータプログラム製品においては、ファイル領域に化合物情報ファイル等が記録され、プログラム領域に反応式検出プログラム

が記録されている。

反応式検出プログラムは、情報処理装置を用いて実行させることができる。この実行によって、まず、第1の処理ルーチンの処理が行われ、入力手段で受け付けられた化合物についてのデータから正準化データが作成される。そして、作成された正準化データに基づいて化合物情報ファイルが検索され、正準化データが化合物情報ファイルに存在する場合には、正準化データに対応する化合物の番号が読み出される。

次に、第2の処理ルーチンの処理が行われ、第1の処理ルーチンで読み出された化合物の番号に基づいて、この化合物が基質あるいは生成物となる酵素の番号が、相互関連情報ファイルから読み出される。さらに、第3の処理ルーチンの処理が行われ、第2の処理ルーチンで読み出された番号の酵素と前記化合物と共に反応系を構成する他の化合物の番号、及び前記酵素についての付加情報が酵素情報ファイルから読み出される。そして、第4の処理ルーチンの処理が行われ、第1の処理ルーチンで読み出された化合物の番号と、第2の処理ルーチンで読み出された酵素の番号と、前記第3の処理部で読み出された他の化合物の番号とから、前記入力手段で受け付けられた化合物の反応式図を表示手段に表示させる。さらに、第3の処理ルーチンで読み出された酵素についての付加情報も表示手段に表示させる。

また、本発明の生化学情報コンピュータプログラム製品は、前記ファイル領域において、受容体の番号と当該受容体の作動物質および／または拮抗物質である化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された、コンピュータ読取可能な受容体情報ファイルを更に有していてもよく、この場合；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係

を示す一覧表を格納しており；

前記コンピュータプログラム製品が、前記プログラム領域において、

化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物が作動物質および／または拮抗物質となる受容体についての付加情報を検出する、コンピュータ読取可能な受容体情報検出プログラムを更に有しており；

前記受容体情報検出プログラムは、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す、コンピュータ読取可能な第5の処理ルーチンと、

前記第5の処理ルーチンで読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す、コンピュータ読取可能な第6の処理ルーチンと、

前記第6の処理ルーチンで読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す、コンピュータ読取可能な第7の処理ルーチンと、

前記第7の処理ルーチンで読み出された受容体についての少なくとも付加情報を前記表示手段に表示する、コンピュータ読取可能な第8の処理ルーチンとを備える。

この場合、上記本発明の生化学情報コンピュータプログラム製品においては、プログラム領域に前記反応式検出プログラムに加えて受容体情報検出プログラムが記録されている。

受容体情報検出プログラムは、情報処理装置を用いて実行させることができる。この実行によって、まず、第5の処理ルーチンの処理が行われ、入力手段で受け

付けられた化合物についてのデータから正準化データが作成される。そして、作成された正準化データに基づいて化合物情報ファイルが検索され、正準化データが化合物情報ファイルに存在する場合には、正準化データに対応する化合物の番号が読み出される。

次に、第6の処理ルーチンの処理が行われ、第5の処理ルーチンで読み出された化合物の番号に基づいて、この化合物が作動物質あるいは拮抗物質となる受容体の番号が、相互関連情報ファイルから読み出される。さらに、第7の処理ルーチンの処理が行われ、第6の処理ルーチンで読み出された番号の受容体についての少なくとも付加情報が受容体情報ファイルから読み出される。そして、第8の処理ルーチンの処理が行われ、第7の処理ルーチンで読み出された受容体についての少なくとも付加情報を表示手段に表示させる。

また、本発明の生化学情報コンピュータプログラム製品は、前記プログラム領域において、反応経路を構成する複数の化合物の中から選択される所定の化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記複数の化合物の反応経路を検出する、コンピュータ読取可能な反応経路検出プログラムを更に有していてもよく、この場合；

前記反応経路検出プログラムは、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す、コンピュータ読取可能な第9の処理ルーチンと、

前記第9の処理ルーチンで読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す、コンピュータ読取可能な第10の処理ルーチンと、
前記第10の処理ルーチンで読み出された酵素の番号に基づいて、この酵素の

基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報ファイルから読み出す、コンピュータ読取可能な第 11 の処理ルーチンと、

前記第 10 の処理ルーチンによる処理と前記第 11 の処理ルーチンによる処理とを繰り返し、所定反応経路内の化合物及び酵素を検索する、コンピュータ読取可能な第 12 の処理ルーチンと、

前記第 10 の処理ルーチンで読み出された酵素の番号と、前記第 11 の処理ルーチンで読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前記表示手段に表示する、コンピュータ読取可能な第 13 の処理ルーチンとを備える。

この場合、上記本発明の生化学情報コンピュータプログラム製品においては、プログラム領域に前記反応式検出プログラム及び前記受容体情報検出プログラムに加えて反応経路検出プログラムが記録されている。

反応経路検出プログラムは、情報処理装置を用いて実行させることができる。この実行によって、まず、第 9 の処理ルーチンの処理が行われ、入力手段で受け付けられた所定の化合物についてのデータから正準化データが作成される。そして、作成された正準化データに基づいて化合物情報ファイルが検索され、正準化データが化合物情報ファイルに存在する場合には、正準化データに対応する化合物の番号が読み出される。

次に、第 10 の処理ルーチンの処理が行われ、第 9 の処理ルーチンで読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とが、相互関連情報ファイルから読み出される。さらに、第 11 の処理ルーチンの処理が行われ、第 10 の処理ルーチンで読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号とが、酵素情報ファイルから読み出される。第 10 の処理ルーチンと第 11 の処理ルーチンとの各処理は、第 12 の処理ルーチンによって繰り返される。

そして、第 13 の処理ルーチンの処理が行われ、第 10 の処理ルーチンで読み出された酵素の番号と、第 11 の処理ルーチンで読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って表示手段に表示させる。

更に、本発明の生化学情報コンピュータプログラム製品は以下のものであってもよい。すなわち、生化学情報を表す画像データまたは生化学情報を表す記号データの入力を受け付ける入力手段と、少なくとも化学反応式の反応式図を表示する表示手段と、コンピュータ使用可能な媒体から情報を読み出す読出手段とを備える情報処理装置と共に用いるための生化学情報コンピュータプログラム製品であって；

ファイルを記録するためのファイル領域とプログラムを記録するためのプログラム領域とを有しており、かつ、前記入力手段から入力されたデータに基づいて少なくとも反応式図を効率的に検索させかつ前記表示手段で表示させるための該媒体中に具現化されたコンピュータ読取可能なファイル及びプログラムを有しているコンピュータ使用可能な媒体を前記コンピュータプログラム製品は備えており；

該コンピュータプログラム製品は、

前記ファイル領域において、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及び当該化合物についての付加情報が格納された、コンピュータ読取可能な化合物情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物である化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納された、コンピュータ読取可能な酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化合物が生成物となる酵素の番号との関係を示す一覧表が格納された、コンピュータ読取可能な相互関連情報ファイルとを有しており、

前記プログラム領域において、

反応経路を構成する複数の化合物の中から選択される所定の化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記複数の化合物の反応経路を検出する、コンピュータ読取可能な反応経路検出プログラムを有しており；

前記反応経路検出プログラムは、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す、コンピュータ読取可能な第 9 の処理ルーチンと、

前記第 9 の処理ルーチンで読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す、コンピュータ読取可能な第 10 の処理ルーチンと、

前記第 10 の処理ルーチンで読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報ファイルから読み出す、コンピュータ読取可能な第 11 の処理ルーチンと、

前記第 10 の処理ルーチンによる処理と前記第 11 の処理ルーチンによる処理とを繰り返し、所定反応経路内の化合物及び酵素を検索する、コンピュータ読取可能な第 12 の処理ルーチンと、

前記第 10 の処理ルーチンで読み出された酵素の番号と、前記第 11 の処理ルーチンで読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前記表示手段に表示する、コンピュータ読取可能な第 13 の処理ルーチンとを備える。

この場合、本発明の生化学情報コンピュータプログラム製品は、前記ファイル領域において、受容体の番号と当該受容体の作動物質および／または拮抗物質で

ある化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された、コンピュータ読取可能な受容体情報ファイルを更に有していてもよく、この場合；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表を格納しており；

前記コンピュータプログラム製品が、前記プログラム領域において、

化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物が作動物質および／または拮抗物質となる受容体についての付加情報を検出する、コンピュータ読取可能な受容体情報検出プログラムを更に有しており；

前記受容体情報検出プログラムは、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す、コンピュータ読取可能な第5の処理ルーチンと、

前記第5の処理ルーチンで読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す、コンピュータ読取可能な第6の処理ルーチンと、

前記第6の処理ルーチンで読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す、コンピュータ読取可能な第7の処理ルーチンと、

前記第7の処理ルーチンで読み出された受容体についての少なくとも付加情報を前記表示手段に表示する、コンピュータ読取可能な第8の処理ルーチンとを備

える。

更に、本発明の生化学情報コンピュータプログラム製品においては、

前記入力手段が、化合物を構成する各原子についての固有データ及び原子間の結合対データの入力を受け付けるものであり；

前記コンピュータプログラム製品が、前記プログラム領域において、前記入力手段で受け付けられた各データに基づいて、前記化合物の化学構造を一意的に特定できる正準化データを作成する、コンピュータ読取可能な正準化データ作成プログラムを更に有していることが好ましい。すなわち、前記正準化データ作成プログラムは、

各原子を等価原子ごとに別のクラスに分類して、クラスごとに異なるクラス番号を各原子に与える、コンピュータ読取可能な構成原子分類ルーチンと、

前記構成原子分類ルーチンで各原子に与えられたクラス番号に基づいて、前記化合物の構造と一意的に対応した正準化番号を各原子に与える、コンピュータ読取可能な正準化番号付与ルーチンと、

前記正準化番号付与ルーチンで各原子に与えられた正準化番号に基づいて、前記正準化データを作成する、コンピュータ読取可能な正準化データ作成ルーチンとを備える。

このような構成を有する本発明にかかる生化学情報コンピュータプログラム製品を所定の情報処理装置に収容して、プログラム領域に格納された正準化データ作成プログラムを読み取ることにより、正準化データ作成プログラムを情報処理装置で実行させることができる。正準化データ作成プログラムの起動によって、まず構成原子分類ルーチンが実行されて、各原子についての固有データ及び原子間の結合対データに基づいて、各原子が等価原子ごとに別のクラスに分類される。そして、クラスごとに異なるクラス番号が各原子に与えられる。次に、正準化番号付与ルーチンが実行されて、各原子に与えられたクラス番号及び原子間の結合対データに基づいて、化合物の構造と一意的に対応した正準化番号が各原子に与

えられる。さらに、正準化データ作成ルーチンが実行されて、各原子に与えられた正準化番号及び各原子についての固有データに基づいて正準化データが作成される。

ここで、前記構成原子分類ルーチンは、

各原子に3種類の属性 (a_i , b_{ij} , d_{ij}) を与え、これらの属性が一つでも異なる原子は非等価であると判定できることを利用して、各原子を等価原子毎に異なるクラス番号を付与しており、

前記3種類の属性 (a_i , b_{ij} , d_{ij}) の中で、 a_i は入力番号 i の原子の種類番号であり、 b_{ij} は入力番号 i の原子に隣接する結合のうち、その種類番号が j である結合の数であり、 d_{ij} は入力番号 i の原子から最短経路により j 個の結合を経て巡れる道筋の数であり；

前記正準化番号付与ルーチンは、

正準化番号を1から昇順に各原子に与える過程において、前記クラス番号の優先順位が最高である原子に正準化番号1を与え、以降正準化番号 n までが付与されている時、既に正準化番号が与えられている原子でかつ未だ正準化番号が与えられていない原子が結合している原子の中で正準化番号が最小である原子を選び、その原子に結合している未だ正準化番号が与えられていない原子の中で前記クラス番号の優先順位が最高である原子に正準化番号 $n+1$ を与えており；

前記正準化データ作成ルーチンは、

各原子に3種類の属性 (P_i , T_i , S_i) を与えて、これらの属性を一行に並べることによって前記正準化データを作成しており、

前記3種類の属性 (P_i , T_i , S_i) の中で、 P_i は正準化番号 i の原子に結合し且つ正準化番号が最小の原子の正準化番号であり、 T_i は正準化番号 i の原子と正準化番号 P_i の原子との結合の種類記号であり、 S_i は正準化番号 i の原子の種類記号であることが好ましい。

なお、本発明にかかる前記コンピュータ使用可能な媒体としては、円盤型記録

媒体、または、テープ型記録媒体が好ましい。

図面の簡単な説明

図 1 は本発明の生化学情報処理装置の一例の構成を示すブロック図である。

図 2 は化合物番号 C₁ の化合物が化合物番号 C₇ の化合物に変化するまでの経路を示す反応経路図の一例である。

図 3 は化合物情報ファイルの構成を示す図である。

図 4 は酵素情報ファイルの構成を示す図である。

図 5 は受容体情報ファイルの構成を示す図である。

図 6 は本発明にかかる相互関連情報ファイルの一例の構成を示す図である。

図 7 は生化学情報処理装置内でのデータの流れを示す図である。

図 8 A は画像データ、図 8 B は結合表データ、図 8 C は正準化データ的具体例をそれぞれ示す図である。

図 9 A は画像データ、図 9 B は結合表データ、図 9 C は正準化データ的具体例をそれぞれ示す図である。

図 10 A ~ C は画像データと正準化データの関係を示す図である。

図 11 はメインルーチンの処理の流れを示すフローチャートである。

図 12 は 3 次元表示ルーチンの処理の流れを示すフローチャートである。

図 13 は反応式検出ルーチンの処理の流れを示すフローチャートである。

図 14 は反応経路検出ルーチンの処理の流れを示すフローチャートである。

図 15 は反応経路検出ルーチンの処理の流れを示すフローチャートである。

図 16 はディスプレイへの表示例を示す図である。

図 17 はディスプレイへの表示例を示す図である。

図 18 A は結合表の原子テーブルの内容を示す図であり、図 18 B は結合表の原子対テーブルの内容を示す図である。

図 19 は正準化データ作成装置の動作の概要を示す概略図である。

図20はメインルーチンの処理の概要を示すフローチャートである。

図21は構成原子分類ルーチンの処理の概要を示すフローチャートである。

図22Aは結合表の原子テーブルの内容を示す図であり、図22Bは結合表の原子対テーブルの内容を示す図である。

図23は3, 5-ジメチル-2, 3, 4, 5-テトラヒドロピリジンを構成する各原子と入力番号との関係を示す図である。

図24A~Bはそれぞれ参照テーブルのデータ内容を示す図である。

図25は3, 5-ジメチル-2, 3, 4, 5-テトラヒドロピリジンを構成する各原子に与えられた3種類の属性(a_i , b_{ij} , d_{ij})を示す図である。

図26A~Bはそれぞれ参照テーブルのデータ内容を示す図である。

図27は参照テーブルのデータ内容を示す図である。

図28A~Bはそれぞれ参照テーブルのデータ内容を示す図である。

図29A~Bはそれぞれ参照テーブルのデータ内容を示す図である。

図30A~Cはそれぞれ3, 5-ジメチル-2, 3, 4, 5-テトラヒドロピリジンを構成する各原子とクラス番号との関係を示す図である。

図31は3, 5-ジメチル-2, 3, 4, 5-テトラヒドロピリジンを構成する各原子に与えられた属性 V_{ij}^1 を示す図である。

図32は3, 5-ジメチル-2, 3, 4, 5-テトラヒドロピリジンを構成する各原子に与えられた属性 V_{ij}^2 を示す図である。

図33は正準化番号付与ルーチンの処理の概要を示すフローチャートである。

図34は3, 5-ジメチル-2, 3, 4, 5-テトラヒドロピリジンを構成する各原子と正準化番号との関係を示す図である。

図35は正準化データ作成ルーチンの処理の概要を示すフローチャートである。

図36Aは結合表の原子テーブルの内容を示す図であり、図36Bは結合表の原子対テーブルの内容を示す図である。

図37は正準化木構造データのデータ内容を示す図である。

図38AはC₆₀の分子構造図であり、図38Bはその正準化データである。

図39は本発明の生化学情報処理装置の他の例の構成を示すブロック図である。

図40は本発明にかかる正準化データ作成装置の一例の構成を示すブロック図である。

図41は本発明の生化学情報処理装置の更に他の例の構成を示すブロック図である。

図42は本発明にかかる相互関連情報ファイルの他の例の構成を示す図である。

図43はメインルーチンの他の例の処理の流れを示すフローチャートである。

図44は本発明の生化学情報記憶媒体の一例の構成を示すブロック図である。

図45は本発明にかかる生化学情報処理装置の一例の構成を示すブロック図である。

図46は本発明にかかる生化学情報処理装置の一例を示す斜視図である。

図47は本発明の生化学情報記憶媒体の他の例の構成を示すブロック図である。

図48は本発明にかかる正準化データ作成用記録媒体の一例の構成を示すブロック図である。

図49は本発明にかかる正準化データ作成装置の他の例の構成を示すブロック図である。

図50は本発明の生化学情報記憶媒体の更に他の例の構成を示すブロック図である。

発明を実施するための最良の形態

以下、本発明の好適な実施例について添付図面を参照して説明する。図1は、本発明の実施例に係る生化学情報処理装置1の構成を示すブロック図である。同図より、本実施例の生化学情報処理装置1は、化合物の分子構造図等を示す画像データを記憶する画像メモリ10と、データを一時的に記憶する作業用メモリ11と、オペレーティングシステム(OS)21及び生化学情報処理プログラム2

2が記憶された第1記憶装置20と、各種ファイルが記憶された記憶手段である第2記憶装置30とを備えている。また、表示手段であるディスプレイ40と、画像データの入力を受け付けるマウス51及び記号データの入力を受け付けるキーボード52を有する入力手段である入力装置50と、画像データ等を出力するプリンタ60と、生化学情報処理プログラム22の実行等を制御するCPU70とを備えている。

生化学情報処理プログラム22は、処理を統括するメインプログラム23と、画像データを3次元表示させる3次元表示プログラム24と、反応式検出手段である反応式検出プログラム25と、受容体情報検出手段である受容体情報検出プログラム26と、反応経路検出手段である反応経路検出プログラム27とを備えている。反応式検出プログラム25は、基質および/または生成物である化合物に関する化学反応式を検出するプログラムで、第1の処理ルーチン25a~第4の処理ルーチン25dを備えている。また、受容体情報検出プログラム26は、受容体についての付加情報を検出するプログラムで、第5の処理ルーチン26a~第8の処理ルーチン26dを備えている。さらに、反応経路検出プログラム27は、複数の化合物の反応経路を検出するプログラムで、第9の処理ルーチン27a~第13の処理ルーチン27eを備えている。

なお、受容体情報検出プログラム26では、ホルモン受容体等の生体本来のもののみならず、薬物等の受容体や存在が確認されていない概念上の受容体も取り扱うことができる。

第2記憶装置30は、化合物情報ファイル31と、酵素情報ファイル32と、相互関連情報ファイル33と、部分相互関連データファイル34と、結合表ファイル（結合表情報ファイルともいう）35と、受容体情報ファイル36とを備えている。この内、化合物情報ファイル31には、化合物番号とこの化合物に対応する正準化データ（カノニカルデータともいう）との関係を示す一覧表、及びこの化合物についての付加情報（例えば図3の参照データ）が格納されている。ま

た、酵素情報ファイル 3 2 には、酵素番号とこの酵素の基質である化合物の化合物番号とこの酵素による生成物である化合物の化合物番号との関係を示す一覧表、及びこの酵素についての付加情報（例えば図 4 の参照データ）が格納されている。さらに、相互関連情報ファイル 3 3 には、化合物番号と当該化合物が基質となる酵素の酵素番号と当該化合物が生成物となる酵素の酵素番号と、当該化合物が作動物質となる受容体の番号と当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表が格納されている。さらにまた、部分相互関連データファイル 3 4 には反応経路情報が、結合表ファイル 3 5 には結合表データがそれぞれ格納できるように用意されている。さらにまた、受容体情報ファイル 3 6 には、受容体番号とこの受容体の作動物質である化合物の化合物番号とこの受容体の拮抗物質である化合物の化合物番号との関係を示す一覧表、及びこの受容体についての付加情報（例えば図 5 の参照データ）が格納されている。

次に、化合物情報ファイル 3 1、酵素情報ファイル 3 2、相互関連情報ファイル 3 3、及び受容体情報ファイル 3 6 の詳細な構成について説明する。図 2 は、化合物番号 C_1 の化合物が、酵素番号 $E_1 \sim E_6$ の複数の酵素を触媒として、化合物番号 C_2, C_3, \dots の化合物へと順次変化し、最終的に化合物番号 C_7 の化合物に変化するまでの経路を示した反応経路図の例であると共に、化合物 $C_6 \sim C_{12}$ が受容体 $R_1 \sim R_4$ に対する作動物質、或いは拮抗物質となっている状況を示す図の例である。

この反応経路図の一例に記載された化合物番号 $C_1 \sim C_7$ は、図 3 に示す化合物情報ファイル 3 1 に記録されている。化合物情報ファイル 3 1 には、化合物番号 $C_1 \sim C_7$ の各化合物に対応した正準化データと、化合物 $C_1 \sim C_7$ の各化合物についての参照データ（名称、文献、物性等）とが、化合物番号 $C_1 \sim C_7$ に対応した一覧表として記録されている。このため、化合物番号 $C_1 \sim C_7$ をキーとして、化合物情報ファイル 3 1 にアクセスすれば、化合物番号 $C_1 \sim C_7$ の各化合物についての、正準化データ及び参照データを読み出すことができる。こ

で、正準化データは各化合物の化学構造を一意的に特定する複数の記号データである。正準化データの詳細は後述する。

また、図2の反応経路図の一例に記載された酵素番号E₁～E₆は、図4に示す酵素情報ファイル32に記録されている。酵素情報ファイル32には、酵素番号E₁～E₆の各酵素の基質である化合物の化合物番号C₁～C₆と、酵素番号E₁～E₆の各酵素による生成物である化合物の化合物番号C₇～C₁₂と、酵素番号E₁～E₆の各酵素についての参照データ（名称、文献、物性、阻害物質、誘導物質、活性化物質等）とが、酵素番号E₁～E₆に対応した一覧表として記録されている。

このため、酵素番号E₁～E₆をキーとして、酵素情報ファイル32にアクセスすれば、酵素番号E₁～E₆の各酵素についての、基質及び生成物である化合物番号C₁～C₁₂と、参照データとを読み出すことができる。また、酵素分類や酵素の同定がなされていない酵素による反応や光、熱、酸、塩基、金属イオン等が関与する非酵素的な反応に加えて、複数の酵素による多段階の反応についても同様に取り扱うことができる。

さらに、受容体番号R₁～R₄は、図5に示す受容体情報ファイル36に記録されている。受容体情報ファイル36には、受容体番号R₁～R₄の各受容体の作動物質である化合物の化合物番号C₆、C₁₀～C₁₂と、受容体番号R₁～R₄の各受容体の拮抗物質である化合物の化合物番号C₇～C₉と、受容体番号R₁～R₄の各受容体についての参照データ（名称、文献、物性、作用等）とが、受容体番号R₁～R₄に対応した一覧表として記録されている。

このため、受容体番号R₁～R₄をキーとして、受容体情報ファイル36にアクセスすれば、受容体番号R₁～R₄の各受容体についての、作動物質及び拮抗物質である化合物番号C₆～C₁₂と、参照データとを読み出すことができる。

さらにまた、化合物番号C₁～C₁₂、酵素番号E₁～E₆、及び受容体番号R₁～R₄の相互の関係は、図6に示す相互関連情報ファイル33に記録されてい

る。詳細には、化合物番号 $C_1 \sim C_6$ の各化合物が基質となる酵素の酵素番号 $E_1 \sim E_6$ と、化合物番号 $C_2 \sim C_7$ の各化合物が生成物となる酵素の酵素番号 $E_1 \sim E_6$ と、化合物番号 C_6 の化合物が阻害する酵素の酵素番号 E_4 とが、化合物番号 $C_1 \sim C_7$ に対応した一覧表として記録されている。また、化合物番号 $C_6, C_{10} \sim C_{12}$ の各化合物が作動物質となる受容体の受容体番号 $R_1 \sim R_4$ と、化合物番号 $C_7 \sim C_9$ の各化合物が拮抗物質となる受容体の受容体番号 R_2, R_4 とが、化合物番号 $C_6 \sim C_{12}$ に対応した一覧表として記録されている。

このため、化合物番号 $C_1 \sim C_7$ をキーとして、相互関連情報ファイル 33 にアクセスすれば、化合物番号 $C_1 \sim C_7$ の各化合物が基質或いは生成物となる酵素の酵素番号 $E_1 \sim E_6$ と、化合物番号 C_6 の化合物が阻害する酵素の酵素番号 E_4 とを読み出すことができる。また、化合物番号 $C_6 \sim C_{12}$ をキーとして、相互関連情報ファイル 33 にアクセスすれば、化合物番号 $C_6 \sim C_{12}$ の各化合物が作動物質或いは拮抗物質となる受容体の受容体番号 $R_1 \sim R_4$ を読み出すことができる。

次に、酵素情報ファイル 32 のデータ内容について、具体的に説明する。まず、図 2 の反応経路図より、酵素番号 E_1 の酵素の基質である化合物の化合物番号は C_1 である。また、酵素番号 E_1 の酵素による生成物である化合物の化合物番号は C_2 である。このため、図 4 の酵素情報ファイル 32 には、酵素番号 E_1 に対応した（基質）化合物番号の欄に、 C_1 が記録されている。また、酵素番号 E_1 に対応した（生成物）化合物番号の欄に、 C_2 が記録されている。

同様に、図 2 の反応経路図より、酵素番号 E_2 の酵素の基質である化合物の化合物番号は C_2 である。また、酵素番号 E_2 の酵素による生成物である化合物の化合物番号は C_3 である。このため、図 4 の酵素情報ファイル 32 には、酵素番号 E_2 に対応した（基質）化合物番号の欄に、 C_2 が記録されている。また、酵素番号 E_2 に対応した（生成物）化合物番号の欄に、 C_3 が記録されている。

このような関係は、酵素番号 $E_3 \sim E_6$ についても同様であり、酵素番号 E_3

～E₆に対応した（基質）化合物番号及び（生成物）化合物番号の欄には、各々、図2の反応経路図に沿った化合物番号C₃～C₇が記録されている。

次に、受容体情報ファイル36のデータ内容について、具体的に説明する。図5に示すように、受容体番号R₁の受容体の作動物質である化合物の化合物番号C₈が、（作動物質）化合物番号の欄に記録されている。また、受容体番号R₂の受容体の拮抗物質である化合物の化合物番号C₉が、（拮抗物質）化合物番号の欄に記録されている。さらに、受容体番号R₃の受容体の作動物質である化合物の化合物番号C₁₀、C₁₁が、（作動物質）化合物番号の欄に記録されている。さらにまた、受容体番号R₄の受容体の作動物質である化合物の化合物番号C₁₂が、（作動物質）化合物番号の欄に記録され、受容体番号R₄の受容体の拮抗物質である化合物の化合物番号C₇、C₉が、（拮抗物質）化合物番号の欄に記録されている。図2の反応経路図によれば、これらの受容体番号と化合物番号との関係は明らかである。

次に、相互関連情報ファイル33のデータ内容について、具体的に説明する。まず、図2の反応経路図より、化合物番号C₁の化合物が基質となる酵素の酵素番号はE₁である。このため、図6の相互関連情報ファイル33には、化合物番号C₁に対応した（基質）酵素番号の欄に、E₁が記録されている。

同様に、図2の反応経路図より、化合物番号C₂の化合物が基質となる酵素の酵素番号はE₂である。また、化合物番号C₂の化合物が生成物となる酵素の酵素番号はE₁である。このため、図6の相互関連情報ファイル33には、化合物番号C₂に対応した（基質）酵素番号の欄に、E₂が記録されている。また、化合物番号C₂に対応した（生成物）酵素番号の欄に、E₁が記録されている。

このような関係は、化合物番号C₃～C₇についても同様であり、化合物番号C₃～C₇（これらが、相互関連情報ファイル33を用いて検索する際のキーとなる）に対応した（基質）酵素番号及び（生成物）酵素番号の欄には、各々、図2の反応経路図に沿った酵素番号E₂～E₆が記録されている。さらに、化合物

番号C₆の化合物は、酵素番号E₆の基質であり、酵素番号E₆の生成物であると同時に、酵素番号E₄の阻害剤であるので、(阻害)酵素番号の欄にE₄が記録されている。

さらにまた、化合物番号C₆の化合物についての作動物質の受容体番号R₁が(作動)受容体番号の欄に記録されている。また、化合物番号C₇の化合物についての拮抗物質の受容体番号R₄が(拮抗)受容体番号の欄に記録されている。以下同様に、化合物番号C₈～C₁₂の化合物についての作動物質/拮抗物質の受容体番号R₂～R₄が、(作動)受容体番号/(拮抗)受容体番号の各欄に記録されている。

次に、生化学情報処理装置1内でのデータの流れを図7に示す。まず、操作者がマウス51を用いて、ディスプレイ40上に分子構造図を作画すると、この分子構造図は画像データ80として、画像メモリ10に記憶される。この画像データ80は、結合表データ81、正準化データ82、3次元データ83のいずれにも変換することができる。

画像データ80・結合表データ81間の変換は、使用OSに対応したグラフィック・ライブラリを使用して行うことができる。なお、結合表データ81・正準化データ82間の変換アルゴリズムについては後で詳細に説明する。また、結合表データ81・3次元データ83間の変換アルゴリズムは、本発明者による“第13回情報科学討論会要旨集 25頁”に記載されている。

変換された結合表データ81は結合表ファイル35に、正準化データ82は作業用メモリ11に、3次元データ83は画像メモリ10にそれぞれ記憶される。また、操作者がキーボード52を用いて、名称等を表わす記号データ84を入力すると、文字列による検索処理84bが化合物情報ファイル31に対して行われ、該当する化合物の正準化データから、結合表データ81が作成される。この結合表データ81も同様に、画像データ80、3次元データ83のいずれかに変換することができる。一方、酵素名等を表わす記号データ84が入力された場合、文

字列による検索処理 8 4 b は酵素情報ファイル 3 2 に対して行われ、該当する酵素番号が読みだされ、以後の処理に用いることができる。

図 8 A ～ 図 8 C に、画像データ 8 0 a、結合表データ 8 1 a、及び正準化データ 8 2 a の具体例を示す。図 8 A は、化合物 “4-メチルピリジン” の分子構造を示す画像データ 8 0 a である。この画像データ 8 0 a は、図 8 B に示す結合表データ 8 1 a に変換することができる。結合表データ 8 1 a は、原子数、結合数、各原子の座標、各元素の元素名などが記録されたテーブルである。この結合表データ 8 1 a を用いれば、全ての化合物の構造を数値データとして表現することができる。

さらに、結合表データ 8 1 a は、図 8 C に示す正準化データ 8 2 a に変換することができる。正準化データ 8 2 a は、数字及び符号等が並んだ記号列である。図 8 C に示すように、化合物 “4-メチルピリジン” の正準化データ 8 2 a は、“1%1%1-2%3%5%N/6%7/” である。このように、正準化データ 8 2 a は、化合物の構造を非常に短い記号列で表すことができる。このため、この正準化データ 8 2 a を、例えば化合物検索システムに適用すれば、検索速度が向上すると共に、記憶資源を有効に利用することができる。

ところが、上述の結合表データでは、化合物を一意的に特定することが困難であり、化合物検索システムに使用することは適当でない。つまり、図 9 A ～ 図 9 C に示すように、画像データ 8 0 b は、画像データ 8 0 a と同一の化合物を表したデータであるにも関わらず、結合表データ 8 1 a と結合表データ 8 1 b とは、全く異なったものになる。このことから、結合表データから化合物を一意的に特定できないことが判る。これに対して、結合表データ 8 1 b を変換した正準化データ 8 2 b は、正準化データ 8 2 a と同一であり、化合物を一意的に特定することができる。

なお、結合表データ 8 1 a 及び 8 1 b において、各データが記録されたテーブルは、原子の番号から質量までのテーブルと、結合原子対から UP/DOWN ま

でのテーブルとに分けられる。従って、例えば結合表データ81aにおいて、原子の番号(4)と元素名(N)とは対応するが、原子の番号(4)と結合原子対(4 5)、結合種(1)及びUP/DOWN(0)とは対応しない。

特に、図10A～図10Cに示すように、2つの画像データ80c, 80dは、同一の化合物を示す画像データであるにも関わらず、見た目には全く異なる。このような画像データ80c, 80dを変換した正準化データ82cは同一であり、化合物を一意的に特定できることが判る。

このように、正準化データは、化合物を一意的に特定できる点で、結合表データに比べて優れており、このため、本実施例の生化学情報処理装置1の各処理では、主に正準化データが用いられている。

一方、結合表データは座標データを有しているので、化合物の分子構造図をディスプレイ40上に表示するのに有用である。更に、その二次元座標データ(X座標、Y座標)は結合表データの他のデータから演算により求められ得るものである(無論、結合の長さや結合間の角度、ディスプレイ上に表示する際の中心位置などは予め指定しておく必要がある)。

次に、本発明の実施例に係る生化学情報処理方法について説明する。この処理方法には、生化学情報処理装置1が用いられる。まず、OS21の制御の下で、生化学情報処理プログラム22のメインプログラム23が起動される。

図11のフローチャートに示すように、メインプログラム23では、まず、入力方法の選択画面がディスプレイ40に表示される(S100)。この画面表示に従って、操作者がマウス51による入力を選択すると(S101)、ディスプレイ40に分子構造図作画用の画面が表示される。次に、操作者がマウス51を用いて、所定の化合物の構造を示す分子構造図を入力すると、この図形は画像データとして受け付けられ、画像メモリ10に記憶される(S102)。この画像データは、ディスプレイ40にも表示される(S103)。そして、この画像データは、上述した変換アルゴリズムに従って、結合表データに変換される(S1

04)。

また、S100の画面表示に従って、操作者がキーボード52による入力を選択すると(S101)、ディスプレイ40に記号列入力画面が表示される。次に、操作者がキーボードを用いて、所定の化合物を特定する化合物名や化学式などの記号列を入力すると、この入力は受け付けられ(S105)、この記号列で特定される化合物の検索(S106)が、化合物情報ファイル31に対して行われ、該当する化合物の正準化データ82から結合表データ81が作成される(S106b)。そして、結合表データは、上述の二次元座標データに基づき画像データに変換され(S107)、この画像データがディスプレイ40に表示される(S108)。

一方、キーボード52による入力で、酵素名等を表わす記号データ84を入力した場合、文字列による検索(S106)が酵素情報ファイル32に対して行われ、該当する酵素番号が読み出され、同様の処理に供することができる。

S104、及びS108の処理終了後、以下のいずれかの処理の選択画面がディスプレイ40に表示される(S109)。この画面表示に従って、操作者が結合表データの保存処理を選択すると、結合表データは結合表ファイル35に書き込まれる(S111)。そして、結合表ファイル35への書き込みが終了した後に、処理はS109に戻る。また、S109の画面表示に従って、操作者が3次元表示処理を選択すると、3次元表示プログラム24が呼び出される(S112)。3次元表示プログラム24は、化合物の分子構造図を3次元表示させるための処理プログラムである。そして、3次元表示プログラム24の処理が終了した後に、処理はS109に戻る。

さらに、S109の画面表示に従って、操作者が反応式検出処理を選択すると、反応式検出プログラム25が呼び出される(S113)。反応式検出プログラム25は、相互関連情報ファイル33等を検索して、化合物の関わる反応式を検出する処理プログラムである。そして、反応式検出プログラム25の処理が終了し

た後に、処理はS109に戻る。さらにまた、S109の画面表示に従って、操作者が反応経路検出処理を選択すると、反応経路検出プログラム27が呼び出される(S114)。反応経路検出プログラム27は、相互関連情報ファイル33等を検索して、複数の化合物の反応経路を検出する処理プログラムである。そして、反応経路検出プログラム27の処理が終了した後に、処理はS109に戻る。

さらにまた、S109の画面表示に従って、操作者が受容体情報表示処理を選択すると、受容体情報検出プログラム26が呼び出される(S115)。受容体情報検出プログラム26は、相互関連情報ファイル33を検索して、特定の化合物の作動受容体番号及び／又は拮抗受容体番号を読み出す(第6の処理ルーチン26b)と共に、受容体情報ファイル36を検索して、読み出された受容体番号の受容体の参照データを検出し(第7の処理ルーチン26c)、更に検出された参照データを表示する(第8の処理ルーチン26d)処理プログラムである。そして、受容体情報検出プログラム26の処理が終了した後に、処理はS109に戻る。さらにまた、S109の画面表示に従って、操作者が終了処理を選択すると、メインプログラムの全処理が終了する。

次に、図12のフローチャートを用いて、S112で呼び出される3次元表示プログラム24の処理について説明する。この処理は、まず、上述した変換アルゴリズムに従って、結合表データが分子構造図の3次元データに変換される(S120)。そして、この3次元データの回転表示等を行うかについての入力促進画面がディスプレイ40に表示される(S121)。この画面で3次元表示プログラム24の起動を選択すると、使用OSに対応したグラフィック・ライブラリを使用して3次元データは画像データに変換され(S124)、この画像データがディスプレイ40に表示される(S125)。さらに、この画面表示に従って、操作者が配座の変更処理、回転処理、拡大処理、縮小処理のいずれかを選択する(S122)と、一般的な3次元グラフィックスの作成手法によって、これらの処理が行われる(S123)。

次に、図13のフローチャートを用いて、S113で呼び出される反応式検出プログラム25の処理について説明する。この処理は、まず、後述する変換アルゴリズムに従って、結合表データが正準化データに変換される(S130)。そして、検索対象の選択画面がディスプレイ40に表示される(S131)。ここで、操作者が反応式を選択する場合には、前のS102或いはS105で、入力された化合物が基質または生成物のいずれであるかを予め指定しておくのが好ましい。或いは、S130の処理の直前で、化合物についての結合表データと共に、基質であるか生成物であるかの指定の入力を受け付けてもよい。

このような条件の下で、S131の画面表示に従って、操作者が反応式を選択すると(S132)、以下の反応式検出処理が行われる。この処理は、まず、化合物情報ファイル31がアクセスされ、化合物の検索が行われる(S133)。この検索処理は、S130で変換された化合物の正準化データに基づいて行われる。この検索処理で、化合物の正準化データと同一の正準化データが、化合物情報ファイル31内に存在しない場合(S134)には、処理を終了する。また、化合物の正準化データと同一の正準化データが化合物情報ファイル31内に存在する場合には、この正準化データに対応する化合物番号が、化合物情報ファイル31から読み出される。

S133で読み出された化合物番号(キー)に基づいて、化合物が基質あるいは生成物となる(前記指定による)酵素番号が、相互関連情報ファイル33から読み出される(S135)。さらに、S135で読み出された酵素番号に基づいて、この酵素番号に対応した(基質)化合物番号、(生成物)化合物番号及び参照データが酵素情報ファイル32から読み出される(S136)。

このように、S133で読み出された化合物番号と、S135で読み出された酵素番号とから、化合物に関わる反応式図が作成され、この反応式図の画像データがディスプレイ40に表示される。また、S136で読み出された酵素についての参照データがディスプレイ40に表示される(S137)。

なお、反応式図の画像データは、得られた（基質）化合物番号の化合物の分子構造図と、（生成物）化合物番号の化合物の分子構造図とを矢印で結び、該矢印の近傍に酵素の参照データ（特に名称）を配した配置でディスプレイ４０に表示されるのが好ましい。化合物番号から分子構造図への変換は、例えば、化合物番号→結合表データ（結合表ファイルにアクセス）→分子構造図（二次元座標を用いて）の順序で行ってもよい。

ここで、第１の処理ルーチン２５ａがＳ１３０からＳ１３３までの処理を行っており、これらの処理は第１のステップに対応する。また、第２の処理ルーチン２５ｂがＳ１３５の処理を行っており、この処理は第２のステップに対応する。さらに、第３の処理ルーチン２５ｃがＳ１３６の処理を行っており、この処理は第３のステップに対応する。さらにまた、第４の処理ルーチン２５ｄがＳ１３７の処理を行っており、この処理は第４のステップに対応する。

本発明において、第１の処理部、ステップ及び処理ルーチンと、第５の処理部、ステップ及び処理ルーチンと、第９の処理部、ステップ及び処理ルーチンとが、それぞれ同じ処理部、ステップ及び処理ルーチンであってもよい。

次に、Ｓ１３１の画面表示に従って、操作者が分子構造図を選択すると（Ｓ１３２）、以下の分子構造図検出処理が行われる。この処理は、まず、化合物情報ファイル３１がアクセスされ、検出対象の化合物の検索が行われる（Ｓ１３８）。検索処理は、Ｓ１３０で変換された化合物の正準化データに基づいて行われる。この検索処理で、検出対象の正準化データと同一の正準化データが、化合物情報ファイル３１内に存在しない場合（Ｓ１３９）には、処理を終了する。また、検出対象の正準化データと同一の正準化データが化合物情報ファイル３１内に存在する場合には、この正準化データに対応する化合物の化合物番号が、化合物情報ファイル３１から読み出される。

Ｓ１３８で読み出された化合物番号に基づいて、参照データ等が化合物情報ファイル３１及び相互関連情報ファイル３３から読み出される（Ｓ１４０）。この

ように、S 1 3 8で読み出された化合物番号から、検出対象である化合物の分子構造図が作成され、この分子構造図の画像データがディスプレイ 4 0に表示される。また、S 1 4 0で読み出されたこの化合物についての参照データがディスプレイ 4 0に表示される (S 1 4 1)。

次に、図 1 4、図 1 5のフローチャートを用いて、S 1 1 4で呼び出される反応経路検出プログラム 2 7の処理について説明する。この処理は、まず、後述する変換アルゴリズムに従って、中心化合物の結合表データが正準化データに変換され、引き続き、検出される反応経路領域を確定するため、所定の反応ステップ数 (例えば、中心化合物を中心として上流側 3 反応ステップ、下流側 5 反応ステップといったように) の入力を受け付ける (S 1 5 0)。

次に、化合物情報ファイル 3 1がアクセスされ、S 1 5 0で変換した正準化データに基づいて、中心化合物の検索が行われる (S 1 5 1)。この検索処理で、中心化合物の正準化データと同一の正準化データが、化合物情報ファイル 3 1内に存在しない場合 (S 1 5 2) には、処理を終了する。また、中心化合物の正準化データと同一の正準化データが化合物情報ファイル 3 1内に存在する場合には、この正準化データに対応する化合物番号が、化合物情報ファイル 3 1から読み出される。

S 1 5 1で読み出された化合物番号 (キー) に基づいて、この化合物を基質とする酵素の酵素番号とこの化合物を生成物とする酵素の酵素番号とが、相互関連情報ファイル 3 3から読み出される (S 1 5 3)。さらに、S 1 5 3で読み出された酵素番号に基づいて、この酵素の基質である化合物の化合物番号とこの酵素による生成物である化合物の化合物番号とが、酵素情報ファイル 3 2から読み出される (S 1 5 4)。そして、S 1 5 3で読み出された酵素番号と、S 1 5 4で読み出された化合物番号とは、順次、部分相互関連データファイル 3 4に追加される (S 1 5 5)。

S 1 5 3からS 1 5 5までの処理は、S 1 5 4で新たに呼び出された化合物番

号について繰り返され、所定ステップ数の反応経路内の全ての化合物の化合物番号と酵素の酵素番号とが、部分相互関連データファイル 34 に書き込まれる (S 156)。

次に、操作者の指示に従って、反応経路内の所定の酵素が指定されると (S 157)、この酵素の基質である化合物、及びこの酵素による生成物である化合物が、化合物情報ファイル 31 と酵素情報ファイル 32 とから読み出され、これらの化合物及び酵素から反応式データが作成される (S 158)。そして、この反応式データがディスプレイ 40 に表示される (S 159)。さらに、部分相互関連データファイル 34 がアクセスされ、この反応式の全ての隣接反応が求められ、これらの隣接反応を示す矢印がディスプレイ 40 に表示される (S 160)。

このようにディスプレイ 40 に表示された反応式データに基づいて、操作者がいずれかの隣接反応を表示させる選択を行った場合 (S 161)、S 157 の処理に戻って、隣接反応についての反応式データが作成される。

ここで、第 9 の処理ルーチン 27 a が S 150、S 151 の処理を行っており、これらの処理は第 9 のステップに対応する。また、第 10 の処理ルーチン 27 b が S 153 の処理を行っており、この処理は第 10 のステップに対応する。さらに、第 11 の処理ルーチン 27 c が S 154 の処理を行っており、この処理は第 11 のステップに対応する。さらにまた、第 12 の処理ルーチン 27 d が S 156 の処理を行っており、この処理は第 12 のステップに対応する。さらにまた、第 13 の処理ルーチン 27 e が S 157 から S 161 までの処理を行っており、これらの処理は第 13 のステップに対応する。

S 159 及び S 160 の処理によるディスプレイ 40 への表示例を、図 16、図 17 に示す。これらの図より、ディスプレイ 40 には、どちらも反応式データを示す画像データ 80 f、80 g が表示されており、反応式データの両端には隣接反応を示す矢印が付加されている。S 161 で行う隣接反応の選択は、いずれかの矢印の部分のマウス 51 でクリックして行う。ここでは、画像データ 80 f

の左端の矢印がマウス 5 1 でクリックされた場合に、1 ステップ前の反応である画像データ 8 0 g が表示される。このような画面の切り替えによって、反応経路内のいずれの反応式の表示も自在に行うことができる。

次に、本発明に好適な正準化データ作成手段及び方法について説明する。

前述の結合表データ 8 1 ・正準化データ 8 2 間の変換アルゴリズムとしてはいずれも、モルガン・アルゴリズム (H.L.Morgan, J.Chem.Doc., 5(2), 107(1965)) のような既知のものが適用可能であり、また、本発明者による“第 1 3 回情報科学討論会要旨集 2 5 頁”に記載された変換アルゴリズムも適用可能である。しかしながら、上記従来の本発明者による変換アルゴリズムにおいては、原子を等価原子毎に分類する処理過程を経ないモルガン・アルゴリズムより高速に正準化データを求めることができたものの、原子の属性としてその原子から特定の最小距離にある原子の数を用いており、等価原子の判別に厳密性が欠け、得られる正準化データの信頼性が未だ充分ではなかった。従って、本発明においては、以下に詳細に説明する正準化データ作成手段及び方法が特に好ましく採用される。

まず、本発明に好適な正準化データ作成手段について説明する。図 1 に示す本発明の実施形態である生化学情報処理装置 1 は本発明にかかる正準化データ作成手段を備えており、すなわち、分子構造図の画像データを記憶する画像メモリ 1 0 と、記号データなどを一時的に記憶する作業用メモリ 1 1 と、オペレーティングシステム (OS) 2 1 及び正準化データ作成プログラム 9 1 が記憶された第 1 記憶装置 2 0 と、結合表ファイル 3 5 及び化合物情報ファイル 3 1 が記憶された第 2 記憶装置 3 0 とを備えている。

また、生化学情報処理装置 1 は、分子構造図を表示するディスプレイ 4 0 と、手書き図形の入力を受け付けるポインティングデバイスであるマウス 5 1 と、化学式等の記号データの入力を受け付けるキーボード 5 2 と、分子構造図を出力するプリンタ 6 0 と、正準化データ作成プログラム 9 1 の実行等を制御する CPU 7 0 とを備えている。なお、ポインティングデバイスには、マウス 5 1 以外に、

タブレット、ディジタイザ、ライトペンなどがあり、これらのいずれの装置をマウス 51 の代わりに備えてもよい。

正準化データ作成プログラム 21 は、化合物を構成する各原子についての固有データ及び原子間の結合対データに基づいて正準化データを作成するプログラムである。この正準化データ作成プログラム 91 は、処理を統括するメインルーチン 91 a と、化合物を構成する各原子にクラス番号を与える構成原子分類ルーチン（構成原子分類処理部）91 b とを備えている。また、正準化データ作成プログラム 91 は、クラス番号に基づいて各原子に正準化番号を与える正準化番号付与ルーチン（正準化番号付与処理部）91 c と、各原子の正準化番号に基づいて正準化データを作成する正準化データ作成ルーチン（正準化データ作成処理部）91 d とを備えている。

第 2 記憶装置 30 には、複数の結合表 81 を格納できる結合表ファイル 35 が設けられている。結合表 81 には化合物を構成する各原子についての固有データ及び原子間の結合対データが記録されており、正準化データ作成プログラム 91 は結合表 81 を介して、これらのデータにアクセスできる。

図 18 A 及び図 18 B に示すように、結合表 81 は、各原子についての固有データが記録された原子テーブル 81 c と、原子間の結合対データが記録された原子対テーブル 81 d とを備えている。具体的には、原子テーブル 81 c には、入力番号（原子の番号ともいう）、原子の二次元座標（X 座標・Y 座標）、元素名（一般に元素記号が用いられる）、属性、原子数、及び結合数を書き込む欄が設けられており（図 18 A 参照）、原子対テーブル 81 d には、結合原子対データ、結合種（例えば、単結合は 1、二重結合は 2 とする）、及び構造（各原子が分子構造図の環状部に属するか、鎖状部に属するかを区別する欄）を書き込む欄が設けられている（図 18 B 参照）。ここで、入力番号は、化合物を構成する各原子をコンピュータで識別するための番号であり、図 18 A の例では数字であるが、記号であっても良い。また、結合原子対データは、入力番号の組合せとして表現

されるのが良い。

なお、正準化データを作成するためには、上記の原子テーブル 8 1 c 及び原子テーブル 8 1 d の全てのデータが必要であるのではなく、固有データとして原子の番号と元素名、結合対データとして結合原子対データと結合種があれば充分である。

また、第 2 記憶装置 3 0 には、化合物番号とこの化合物に対応する正準化データ（カノニカルデータともいう）との関係を示す一覧表が記録された化合物情報ファイル 3 1 が格納されている。図 3 に示すように、化合物情報ファイル 3 1 は、化合物番号 $C_1 \sim C_7$ の各化合物に対応した正準化データと、化合物 $C_1 \sim C_7$ の各化合物についての参照データ（名称、文献、物性等）とを、化合物番号 $C_1 \sim C_7$ に対応した一覧表として記録したファイルである。このため、化合物番号 $C_1 \sim C_7$ をキーとして、化合物情報ファイル 3 1 にアクセスすれば、化合物番号 $C_1 \sim C_7$ の各化合物についての、正準化データ及び参照データを読み出すことができる。ここで、正準化データは各化合物の化学構造を一意的に特定する複数の記号からなるデータである。

なお、構成原子分類ルーチン 9 1 b が構成原子分類ステップに、正準化番号付与ルーチン 9 1 c が正準化番号付与ステップに、正準化データ作成ルーチン 9 1 d が正準化データ作成ステップにそれぞれ対応する。

次に、正準化データ作成手段の動作の概要について説明する。図 1 9 に示すように、操作者はマウス 5 1 又はキーボード 5 2 を操作して、正準化データの作成対象となる化合物の結合表 8 1 を結合表ファイル 3 5 内に作成することができる。

マウス 5 1 による入力、マウス 5 1 を用いてディスプレイ 4 0 上に化合物の分子構造図を手書き入力するもので、入力順に定まる各原子の入力番号が、第 2 記憶装置 3 0 内に作成された結合表 8 1 の入力番号の欄に書き込まれる。さらに、この分子構造図 E₁ の各原子の結合関係を示す結合原子対データが、結合表 8 1 の結合原子対の欄に書き込まれる。このように、マウス 5 1 による入力では、化

化合物を特定する結合表 8 1 が手書きされた原子構造図 E₁ から作成される。

また、キーボード 5 2 による入力、キーボード 5 2 を用いて所定の化合物に対応する結合表名を特定する記号列を入力するもので、入力された記号データ 1 1 a に基づいて、この結合表名によって特定される結合表 8 1 が結合表ファイル 3 5 から読み出される。

このように、マウス 5 1 及びキーボード 5 2 が入力手段 A (5 0) を構成し、マウス 5 1 とキーボード 5 2 とのいずれを用いても結合表 8 1 が得られる。そして、正準化作成手段 B である正準化データ作成プログラム 9 1 が実行されて、結合表 8 1 の各データに基づいて正準化データ 8 2 が作成される。このように作成された正準化データ 8 2 は、化合物情報ファイル 3 1 に書き込まれて保存される。ここで、結合表 8 1 から正準化データ 8 2 を作成して保存するのは、結合表 8 1 のままで保存するよりも記憶領域を小さくできかつ一意的に化合物が特定されるからである。即ち、図 1 8 A ~ B に示した結合表 8 1 に基づいて作成した正準化データ 8 2 は “1 % 1 % 1 - 2 % 3 % 5 % N / 6 % 7 /” であり、化合物の構造を非常に短い文字・数字・記号列でかつ一意的に表すことができる。このように短い記号列を保存の対象とすれば、記憶資源を有効に利用でき装置の小型軽量化に寄与することができる。

また、結合表 8 1 の各データに基づいて、二次元座標演算処理が行われることによって、各原子の二次元座標データを得ることができる。このように得られた二次元座標データから美的に優れた分子構造図 E₂ が作成される。作成された分子構造図 E₂ は、ディスプレイ 4 0 に表示することやプリンタ 6 0 から出力することができる。

なお、キーボード 5 2 による入力は、原子の結合状態を示す前記データ等を第 2 記憶装置 3 0 内に作成された結合表 8 1 に直接書き込んでも良い。また、イメージスキャナーやオプティカルカードリーダー (OCR) などの光学的に図形や文字を読み取る装置を本発明の入力装置として用いて、結合表データの入力の受

付けを行っても良い。

次に、本発明に係る実施形態である正準化データ作成方法について説明する。この作成方法には、上述の正準化データ作成手段が用いられる。まず、OS 21の制御の下で、正準化データ作成プログラム91のメインルーチン91aが起動される。

図20のフローチャートに示すように、メインルーチン91aは、まず、構成原子分類ルーチン91bを呼び出して、化合物を構成する各原子にクラス番号を付与する(S910)。次に、正準化番号付与ルーチン91cを呼び出して、各原子に付与されたクラス番号に基づいて各原子に正準化番号を付与する(S920)。さらに、正準化データ作成ルーチン91dを呼び出して、各原子に付与された正準化番号に基づいて正準化データを作成する(S930)。このように作成された正準化データは、化合物情報ファイル31に書き込まれて保存される。

次に、S910で呼び出される構成原子分類ルーチン91bの処理について説明する。この処理は、化合物を構成する各原子を等価原子毎に別のクラスに分類して、属するクラスに対応するクラス番号を各原子に与える処理である。例えば、ベンゼンの各原子は全て等価なので、全てに同一のクラス番号が与えられる。また、トルエンの各原子はそれぞれ非等価なので、それぞれ異なったクラス番号が与えられる。

図21のフローチャートに示すように、まず、結合表81に基づいて化合物を構成する各原子に3種類の属性(a_i , b_{ij} , d_{ij})をそれぞれ与える(S911)。ここで、属性 a_i は入力番号iの原子の種類番号(この例では原子番号)である。また、属性 b_{ij} は入力番号iの原子に隣接する結合のうち、その種類番号(この例では結合種(単結合は1、二重結合は2、三重結合は3、芳香結合は4、…))がjである結合の数(ベクトル量)である。さらに、属性 d_{ij} は入力番号iの原子から最短経路によりj個の結合を経て巡れる道筋の数(ベクトル量)である。

次に、原子ごとに属性 (a_i , b_{ij} , d_{ij}) を並べて9桁の数字列とし、この数字列が小さい順番にクラス番号 C_i^0 を与えて、各原子を複数のクラスに分類する (S 9 1 2)。ここで与えられるクラス番号 C_i^0 は0次のクラス番号であり、S 9 1 3以降のループ処理で1次のクラス番号 C_i^1 、2次のクラス番号 C_i^2 、...を順次求めていく。

次に、次数 n を1にする (S 9 1 3)。そして、各原子に属性 V_{ij}^1 を与える (S 9 1 4)。属性 V_{ij}^n は入力番号 i の原子に結合し、次数 $n-1$ においてクラス番号が j である原子の数である。さらに、原子ごとに属性 (a_i , b_{ij} , d_{ij} , V_{ij}^n) を並べて、この数字列が小さい順番にクラス番号 C_i^n を与えて、各原子を複数のクラスに分類する (S 9 1 5)。そして、クラスの数 N_n が $N_{(n-1)}$ と等しいか調べて、等しい場合には処理を終了する。または、クラスの数 N_n が総原子数と等しいか調べて、等しい場合には処理を終了する (S 9 1 6)。いずれも等しくない場合には、 n に1を加えて処理を S 9 1 4 に戻す (S 9 1 7)。

次に、3, 5-ジメチル-2, 3, 4, 5-テトラヒドロピリジン为例にして構成原子分類ルーチン 9 1 b のステップごとの処理を詳細に説明する。

まず、S 9 1 1 の処理が実行される。この処理を実行する際には、図 2 2 A ~ B に示すようなデータが既に結合表 8 1 に書き込まれており、この結合表 8 1 に書き込まれた各データに基づいて、各原子に3種類の属性 (a_i , b_{ij} , d_{ij}) を与える。ここで、この結合表 8 1 に記録された入力番号は、図 2 3 に示すように各原子が手書き入力された順番に与えられる任意の数字である。

属性 a_i は次のように求められる。上述したように属性 a_i は入力番号 i の原子の種類番号である。ここで、結合表 8 1 には各原子の元素名が記録されており、これらの元素名から種類番号を求めることができる。従って、元素名を結合表 8 1 から読み出すことによって、この元素名に対応した属性 a_i を得ることができる。その結果、 a_1 , a_2 , $a_4 \sim a_8 = 6$, $a_3 = 7$ がそれぞれ得られる。

また、属性 b_{ij} は次のように求められる。上述したように属性 b_{ij} は入力番号 i の原子に隣接する結合のうち、その種類番号が j である結合の数である。結合表 8 1 には各原子の結合種が記録されており、この結合種を結合表 8 1 から読み出すことによって属性 b_{ij} を得ることができる。その結果、 $b_{1j} = (3, 0, 0, 0)$ 、 $b_{2j} = (1, 1, 0, 0)$ 、 $b_{3j} = (1, 1, 0, 0)$ 、 $b_{4j} = (2, 0, 0, 0)$ 、 $b_{5j} = (3, 0, 0, 0)$ 、 $b_{6j} = (2, 0, 0, 0)$ 、 $b_{7j} = (1, 0, 0, 0)$ 、 $b_{8j} = (1, 0, 0, 0)$ がそれぞれ得られる。

具体的には、属性 b_{ij} は図 2 4 A ~ B に示す参照テーブル T を用いて求められる。参照テーブル T は 2 つの原子の結合関係を示す配列 D (x, y) として構成されており、結合表 8 1 の結合原子対及び結合種のデータに基づいて作成される。即ち、結合原子対が示す配列要素に結合種 j が書き込まれて、参照テーブル T が作成されている。

この参照テーブル T を用いた属性 b_{ij} の抽出は次のように行われる。まず、参照テーブル T の配列要素のうち、 $x = 1$ 又は $y = 1$ を満たす配列要素 (図 2 4 A において斜線で示した配列要素) を探索して、この配列要素に書き込まれたデータ (結合種) j を抽出する。その結果、 $D(1, 2) = 1$ 、 $D(1, 6) = 1$ 、 $D(1, 8) = 1$ が得られる。このように得られた 3 つの配列要素のデータ j は全て 1 なので、 $b_{11} = 3$ となる。また、データ j が 2 以上の配列要素は得られないので、 $b_{12} \sim b_{14} = 0$ となる。

次に、参照テーブル T の配列要素のうち、 $X = 2$ 又は $Y = 2$ を満たす配列要素 (図 2 4 B において斜線で示した配列要素) を探索して、この配列要素に書き込まれたデータを抽出する。その結果、 $D(1, 2) = 1$ 、 $D(2, 3) = 2$ が得られる。このように得られた配列要素のデータ j は 1, 2 で、それぞれ 1 つなので、 $b_{21} = b_{22} = 1$ となる。また、データ j が 3 以上の配列要素は得られないので、 $b_{23} = b_{24} = 0$ となる。

$i = 3 \sim 8$ についても同様に処理することにより、図 2 5 に示す属性 b_{ij} (i

$= 1 \sim 8, j = 1 \sim 4$) が得られる。

さらに、属性 d_{ij} は次のように求められる。上述したように属性 d_{ij} は入力番号 i の原子から最短経路により j 個の結合を経て巡れる道筋の数である。即ち、図 2 3 の分子構造図に基づいて説明すると、入力番号 1 の原子から 1 個の結合を経て巡れる道筋は、(入力番号 1 ~ 入力番号 2)、(入力番号 1 ~ 入力番号 6)、(入力番号 1 ~ 入力番号 8) の計 3 本である。また、入力番号 1 の原子から 2 個の結合を経て巡れる道筋は、(入力番号 1 ~ 入力番号 2 ~ 入力番号 3)、(入力番号 1 ~ 入力番号 6 ~ 入力番号 5) の計 2 本である。

さらに、入力番号 1 の原子から最短経路により 3 個の結合を経て巡れる道筋は、(入力番号 1 ~ 入力番号 2 ~ 入力番号 3 ~ 入力番号 4)、(入力番号 1 ~ 入力番号 6 ~ 入力番号 5 ~ 入力番号 4)、(入力番号 1 ~ 入力番号 6 ~ 入力番号 5 ~ 入力番号 7) の計 3 本である。さらにまた、入力番号 1 の原子から最短経路により 4 個の結合を経て巡れる道筋は 0 本である。以上の処理の結果、 $d_{1j} = (3, 2, 3, 0)$ が得られる。

同様の処理によって、 $d_{2j} = (2, 3, 2, 2)$ 、 $d_{3j} = (2, 2, 4, 0)$ 、 $d_{4j} = (2, 3, 2, 2)$ 、 $d_{5j} = (3, 2, 3, 0)$ 、 $d_{6j} = (2, 4, 2, 0)$ 、 $d_{7j} = (1, 2, 2, 3)$ 、 $d_{8j} = (1, 2, 2, 3)$ がそれぞれ得られる。

具体的には、属性 d_{ij} も属性 b_{ij} と同様に、参照テーブル T を参照して求められる。この参照テーブル T を参照した属性 d_{ij} の抽出は $i = 1$ 、 $i = 2$ 、…の順番で行われる。まず、属性 d_{1j} ($i = 1$) が抽出される。

属性 d_{1j} ($i = 1$) の抽出は、参照テーブル T の配列要素のうち、 $X = 1$ 又は $Y = 1$ を満たす配列要素 (図 2 6 A において斜線で示した配列要素) を探索して、データが書き込まれた配列要素を抽出する。そして、抽出された配列要素に結合経路数として 1 を書き込む。その結果、 $D(1, 2)$ 、 $D(1, 6)$ 、 $D(1, 8)$ に結合経路数 1 が書き込まれる (図 2 6 A では結合経路数を三角形で囲んで

示す)。

次に、結合経路数 1 が書き込まれた各配列要素の添字 $S = (1, 2), (1, 6), (1, 8)$ を抽出する。これらの添字 S から前段の抽出処理で利用した 1 を除き、 $S = 2, 6, 8$ を得る。このようにして得られた $S = 2, 6, 8$ に基づいて、 $X = 2, 6, 8$ 又は $Y = 2, 6, 8$ を満たす配列要素 (図 26B において斜線で示した配列要素) を探索し、データが書き込まれて且つ結合経路数が書き込まれていない配列要素を抽出する。そして、抽出された配列要素に結合経路数として 2 を書き込む。その結果、 $D(2, 3), D(5, 6)$ に結合経路数 2 が書き込まれる。

さらに、結合経路数 2 が書き込まれた各配列要素の添字 $S = (2, 3), (5, 6)$ を抽出する。これらの添字 S から前段の抽出処理で利用した 2, 6 を除き、 $S = 3, 5$ を得る。このようにして得られた $S = 3, 5$ に基づいて、 $X = 3, 5$ 又は $Y = 3, 5$ を満たす配列要素 (図 27 において斜線で示した配列要素) を探索し、データが書き込まれて且つ結合経路数が書き込まれていない配列要素を抽出する。そして、抽出された配列要素に結合経路数として 3 を書き込む。その結果、 $D(3, 4), D(4, 5), D(5, 7)$ に結合経路数 3 が書き込まれる。

以上の処理によって、全ての配列要素に結合経路数が書き込まれる。その結果、結合経路数 1 の配列要素が 3 個、結合経路数 2 の配列要素が 2 個、結合経路数 3 の配列要素が 3 個、結合経路数 4 の配列要素が 0 個となり、 $d_{1j} = (3, 2, 3, 0)$ が得られる。

次に、属性 $d_{2j} (i = 2)$ が抽出される。属性 $d_{2j} (i = 2)$ の抽出は、参照テーブル T の配列要素のうち、 $X = 2$ 又は $Y = 2$ を満たす配列要素 (図 28A において斜線で示した配列要素) を探索して、データが書き込まれた配列要素を抽出する。そして、抽出された配列要素に結合経路数として 1 を書き込む。その結果、 $D(1, 2), D(2, 3)$ に結合経路数 1 が書き込まれる (図 28A では結合経路数を三角形で囲んで示す)。

次に、結合経路数 1 が書き込まれた各配列要素の添字 $S = (1, 2), (2, 3)$ を抽出する。これらの添字 S から前段の抽出処理で利用した 2 を除き、 $S = 1, 3$ を得る。このようにして得られた $S = 1, 3$ に基づいて、 $X = 1, 3$ 又は $Y = 1, 3$ を満たす配列要素（図 28 B において斜線で示した配列要素）を探索し、データが書き込まれて且つ結合経路数が書き込まれていない配列要素を抽出する。そして、抽出された配列要素に結合経路数として 2 を書き込む。その結果、 $D(1, 6), D(1, 8), D(3, 4)$ に結合経路数 2 が書き込まれる。

さらに、結合経路数 2 が書き込まれた各配列要素の添字 $S = (1, 6), (1, 8), (3, 4)$ を抽出する。これらの添字 S から前段の抽出処理で利用した 1, 3 を除き、 $S = 4, 6, 8$ を得る。このようにして得られた $S = 4, 6, 8$ に基づいて、 $X = 4, 6, 8$ 又は $Y = 4, 6, 8$ を満たす配列要素（図 29 A において斜線で示した配列要素）を探索し、データが書き込まれて且つ結合経路数が書き込まれていない配列要素を抽出する。そして、抽出された配列要素に結合経路数として 3 を書き込む。その結果、 $D(4, 5), D(5, 6)$ に結合経路数 3 が書き込まれる。

さらにまた、結合経路数 3 が書き込まれた各配列要素の添字 $S = (4, 5), (5, 6)$ を抽出する。これらの添字 S から前段の抽出処理で利用した 4, 6 を除き、 $S = 5, 5$ （即ち、 $S = 5$ が 2 重に適用される。）を得る。このようにして得られた $S = 5, 5$ に基づいて $X = 5$ 又は $Y = 5$ を満たす配列要素（図 29 B において斜線で示した配列要素）を探索して、データが書き込まれて且つ結合経路数が書き込まれていない配列要素を抽出する。そして、抽出された配列要素に結合経路数として 4 を書き込む。その結果、 $D(5, 7)$ に結合経路数 4 が 2 つ書き込まれる。

以上の処理によって、全ての配列要素に結合経路数が書き込まれる。その結果、結合経路数 1 の配列要素が 2 個、結合経路数 2 の配列要素が 3 個、結合経路数 3 の配列要素が 2 個、結合経路数 4 の配列要素が 2 個となり、 $d_{ij} = (2, 3, 2,$

2) が得られる。

$i = 3 \sim 8$ についても同様に処理することにより、図25に示す d_{ij} ($i = 1 \sim 8$, $j = 1 \sim 4$) が得られる。以上説明したS911の処理によって、3, 5-ジメチル-2, 3, 4, 5-テトラヒドロピリジンを構成する各原子に3種類の属性 (a_i , b_{ij} , d_{ij}) が与えられた。

次に、S912の処理が実行される。上述したように、S912では原子ごとに属性 (a_i , b_{ij} , d_{ij}) を並べて9桁の数字列とし、この数字列が小さい順番にクラス番号 C_i^0 を与えて、各原子を複数のクラスに分類している。ここで与えられるクラス番号 C_i^0 は0次のクラス番号である。

S912の処理を具体的に説明すると、入力番号1の原子の数字列は“630003230”であり、入力番号2の原子の数字列は“611002322”である。以下順番に、“711002240”、“620002322”、“630003230”、“620002420”、“610001223”、“610001223”となる。

その結果、入力番号7及び8の原子の数字列が最小となり、これらの原子にクラス番号 $C_7^0 = C_8^0 = 1$ が与えられる。同様に、入力番号2の原子にクラス番号 $C_2^0 = 2$ が与えられて、入力番号4の原子にクラス番号 $C_4^0 = 3$ が与えられる。また、入力番号6の原子にクラス番号 $C_6^0 = 4$ が与えられて、入力番号1及び5の原子にクラス番号 $C_1^0 = C_5^0 = 5$ が与えられる。さらに、入力番号3の原子にクラス番号 $C_3^0 = 6$ が与えられる (図30A参照)。このようにして各原子は6つのクラスに分類されて、クラスの数 N_0 は6となる。

次に、S913の処理が実行されて、次数 n を1にする。

さらに、S914の処理が実行される。上述したように、S914では各原子に属性 $V_{ij}^{1(n=1)}$ を与えている。ここで、属性 $V_{ij}^{1(n=1)}$ は入力番号 i の原子に結合し、クラス番号が j である原子の数である。即ち、図30Bの分子構造図に基づいて説明すると、入力番号1の原子に結合した原子の入力番号は2, 6, 8であ

り、これらの原子のクラス番号は $C_2^0 = 2$ 、 $C_6^0 = 4$ 、 $C_8^0 = 1$ である。その結果、 $j = 1, 2, 4$ の属性 V_{1j}^1 に 1 が書き込まれて、 $V_{1j}^1 = (1, 1, 0, 1, 0, 0)$ が得られる。

また、入力番号 2 の原子に結合した原子の入力番号は 1, 3 であり、これらの原子のクラス番号は $C_1^0 = 5$ 、 $C_3^0 = 6$ である。その結果、 $j = 5, 6$ の属性 V_{2j}^1 に 1 が書き込まれて、 $V_{2j}^1 = (0, 0, 0, 0, 1, 1)$ が得られる。入力番号 3～8 の原子についても同様に処理することにより、 $V_{3j}^1 = (0, 1, 1, 0, 0, 0)$ 、 $V_{4j}^1 = (0, 0, 0, 1, 1, 0)$ 、 $V_{5j}^1 = (1, 0, 1, 1, 0, 0)$ 、 $V_{6j}^1 = (0, 0, 0, 0, 2, 0)$ 、 $V_{7j}^1 = (0, 0, 0, 0, 1, 0)$ 、 $V_{8j}^1 = (0, 0, 0, 0, 1, 0)$ がそれぞれ得られる。

具体的には、属性 V_{ij}^n は図 2 4 A～B に示す参照テーブル T を用いて求められる。この参照テーブル T を用いた属性 V_{ij}^1 の抽出は $i = 1, i = 2, \dots$ の順番で行われる。まず、属性 V_{1j}^1 ($i = 1$) が抽出される。属性 V_{1j}^1 ($i = 1$) の抽出は、参照テーブル T の配列要素のうち、 $x = 1$ 又は $y = 1$ を満たす配列要素 (図 2 4 A において斜線で示した配列要素) を探索して、データが書き込まれた配列要素の添字 $S = (1, 2), (1, 6), (1, 8)$ を抽出する。これらの添字 S から $i = 1$ を除き、 $S = 2, 6, 8$ を得る。このようにして得られた S の値を 0 次のクラス番号 C_i^0 に代入して、 $C_2^0 = 2$ 、 $C_6^0 = 4$ 、 $C_8^0 = 1$ を得る。そして、 $j = 1, 2, 4$ の属性 V_{1j}^1 に 1 を書き込むことにより、 $V_{1j}^1 = (1, 1, 0, 1, 0, 0)$ が求められる。

次に、属性 V_{2j}^1 ($i = 2$) が抽出される。属性 V_{2j}^1 ($i = 2$) の抽出は、参照テーブル T の配列要素のうち、 $X = 2$ 又は $Y = 2$ を満たす配列要素 (図 2 4 B において斜線で示した配列要素) を探索して、データが書き込まれた配列要素の添字 $S = (1, 2), (2, 3)$ を抽出する。これらの添字 S から $i = 2$ を除き、 $S = 1, 3$ を得る。このようにして得られた S の値を 0 次のクラス番号 C_i^0 に代入して、 $C_1^0 = 5$ 、 $C_3^0 = 6$ を得る。そして、 $j = 5, 6$ の属性 V_{2j}^1

1 に 1 を書き込むことにより、 $V_{2j}^1 = (0, 0, 0, 0, 1, 1)$ が求められる。

$i = 3 \sim 8$ についても同様に処理することにより、図 3 1 に示す属性 V_{ij}^1 ($i = 1 \sim 8, j = 1 \sim 6$) が得られる。

次に、S 9 1 5 の処理が実行される。上述したように、S 9 1 5 では原子ごとに属性 (C_i^{n-1}, V_{ij}^n) を並べて、この数字列が小さい順番にクラス番号 C_i^n を与えて、各原子を複数のクラスに分類している。

具体的には、入力番号 1 の原子の数字列は “5 1 1 0 1 0 0” であり、入力番号 2 の原子の数字列は “2 0 0 0 0 1 1” である。以下順番に、“6 0 1 1 0 0 0”、“3 0 0 0 1 1 0”、“5 1 0 1 1 0 0”、“4 0 0 0 0 2 0”、“1 0 0 0 0 1 0”、“1 0 0 0 0 1 0” となる。

その結果、入力番号 7 及び 8 の原子の数字列が最小となり、これらの原子にクラス番号 $C_7^1 = C_8^1 = 1$ が与えられる。同様に、入力番号 2 の原子にクラス番号 $C_2^1 = 2$ が与えられて、入力番号 4 の原子にクラス番号 $C_4^1 = 3$ が与えられる。さらに、入力番号 6 の原子にクラス番号 $C_6^1 = 4$ が与えられて、入力番号 5 の原子にクラス番号 $C_5^1 = 5$ が与えられる。さらにまた、入力番号 1 の原子にクラス番号 $C_1^1 = 6$ が与えられて、入力番号 3 の原子にクラス番号 $C_3^1 = 7$ が与えられる。このようにして各原子は 7 つのクラスに分類されて、クラスの数 N_1 は 7 となる。

次に、S 9 1 6 の処理を実行して、クラスの数 N_n が $N_{(n-1)}$ と等しいか調べて、等しい場合には処理を終了する。また、クラスの数 N_n が総原子数と等しいか調べて、等しい場合には処理を終了する。ここでは、クラスの数 N_1 が 7 で、クラスの数 N_2 が 6 なので、 N_1 と N_2 とは等しくない。また、総原子数は 8 なので、クラスの数 N_1 と総原子数とは等しくない。このようにいずれも等しくないので、S 9 1 7 の処理を実行して n を 2 とする。

さらに、S 9 1 4 に戻って各原子に属性 V_{ij}^2 を与える。その結果、図 3 2 に

示すように、 $V_{1j}^2 = (1, 1, 0, 1, 0, 0, 0)$ 、 $V_{2j}^2 = (0, 0, 0, 0, 0, 1, 1)$ 、 $V_{3j}^2 = (0, 1, 1, 0, 0, 0, 0)$ 、 $V_{4j}^2 = (0, 0, 0, 0, 1, 0, 1)$ 、 $V_{5j}^2 = (1, 0, 1, 1, 0, 0, 0)$ 、 $V_{6j}^2 = (0, 0, 0, 0, 1, 1, 0)$ 、 $V_{7j}^2 = (0, 0, 0, 0, 1, 0, 0)$ 、 $V_{8j}^2 = (0, 0, 0, 0, 0, 1, 0)$ が得られる。

そして、S 9 1 5 の処理を実行して、各原子にクラス番号 C_i^2 を与える。その結果、図 3 0 C に示すように、 $C_1^2 = 7$ 、 $C_2^2 = 3$ 、 $C_3^2 = 8$ 、 $C_4^2 = 4$ 、 $C_5^2 = 6$ 、 $C_6^2 = 5$ 、 $C_7^2 = 2$ 、 $C_8^2 = 1$ が得られる。このようにして各原子は 8 つのクラスに分類されて、クラスの数 N_2 は 8 となる。クラスの数 $N_2 = 8$ は総原子数と等しいので、S 9 1 6 の判定によって処理を終了する。

次に、図 3 3 のフローチャートを用いて、図 2 0 の S 9 2 0 で呼び出される正準化番号付与ルーチン 9 1 c の処理について説明する。ここで正準化番号とは、化合物の構造によって一意的に定まる各原子の番号である。即ち、分子構造図を手書き入力することによって与えられる入力番号は、入力する順番が異なることによって変わる任意的な番号である。これに対して正準化データ 8 2 は、化合物の構造にのみ依存した一意的なデータである。このため、任意的な入力番号から一意的な正準化データ 8 2 を直接作成することは困難である。そこで、正準化データ作成プログラム 9 1 では、入力番号を一旦正準化番号に変換して、この一意的な正準化番号に基づいて正準化データ 8 2 を作成することにより、円滑な正準化データ 8 2 の作成を可能にしている。

正準化番号付与ルーチン 9 1 c の処理は、まず、変数 k に 1 を与える (S 9 2 1)。次に、構成原子分類ルーチン 9 1 b で得られた最終クラス番号 C_i^1 を調べて、最大の原子に正準化番号 k (ここでは $k = 1$) を与える (S 9 2 2)。最大の原子が複数個ある場合には、これらの原子の中から任意の原子を選び、この原子に正準化番号 k を与える。そして、全ての原子に正準化番号が与えられた後に処理を終了する (S 9 2 3)。

次に、変数 k に 1 を加えて (S 9 2 4)、正準化番号の決まった原子 (以下、既決原子という) の中から、まだ正準化番号が決まっていない原子 (以下、未決原子という) が結合している既決原子を抽出する (S 9 2 5)。そして、抽出された既決原子が複数あるか判定して (S 9 2 6)、抽出された既決原子が複数ある場合には、これらの既決原子の中で正準化番号が最小の既決原子を選択する (S 9 2 7)。そして、選択された既決原子に結合している未決原子の中からクラス番号 C_i^r が最大の未決原子を抽出して、この未決原子の正準化番号を k とする (S 9 2 8)。なお、クラス番号 C_i^r が最大の既決原子が複数ある場合には、これらの既決原子の中から任意に選択する。

また、S 9 2 6 で既決原子が 1 つであると判定された場合には、この既決原子に結合している未決原子の中からクラス番号 C_i^r が最大の未決原子を選択して、この未決原子に正準化番号 k を与える (S 9 2 9)。S 9 2 8 及び S 9 2 9 の処理が終了した後に S 9 2 3 に処理を戻し、全ての原子に正準化番号が与えられるまで、S 9 2 3 ~ S 9 2 9 のループを繰り返す。

次に、正準化番号付与ルーチン 9 1 c の処理について、3, 5-ジメチル-2, 3, 4, 5-テトラヒドロピリジンを用いた具体例を説明する。まず、S 9 2 1 の処理で変数 k に 1 を与えて、次に S 9 2 2 の処理が行う。S 9 2 2 の処理では、入力番号 3 の原子が $C_3^r = 8$ で最大なので、入力番号 3 の原子に正準化番号 $k = 1$ を与える。次に、S 9 2 4 の処理で変数 k を 2 にして、S 9 2 5 の処理で入力番号 3 の原子を既決原子として抽出する。

このように抽出された既決原子は 1 つだけなので、次に S 9 2 9 の処理を行う。入力番号 3 の原子に結合している未決原子は入力番号 2, 4 の原子なので、これらの原子の中からクラス番号 C_i^r が最大の原子を選択する。即ち、入力番号 2 の原子のクラス番号は $C_2^r = 3$ で、入力番号 4 の原子のクラス番号は $C_4^r = 4$ である。このため、入力番号 4 の原子を選択して、この原子に正準化番号 $k = 2$ を与える。

次に、S 9 2 4 の処理に戻って変数 k を 3 にして、S 9 2 5 の処理で入力番号 3, 4 の原子を既決原子として抽出する。このように抽出された既決原子は複数あるので、次に S 9 2 7 の処理を行い、抽出された既決原子の中から正準化番号が最小の原子を選択する。即ち、入力番号 3 の原子の正準化番号は 1 で、入力番号 4 の原子の正準化番号は 2 である。このため、入力番号 3 の原子を選択する。そして、S 9 2 8 の処理を行い、入力番号 3 の原子に結合した入力番号 2 の原子に正準化番号 $k = 3$ を与える。

さらに、S 9 2 4 の処理に戻って変数 k を 4 にして、S 9 2 5 の処理で入力番号 2, 4 の原子を既決原子として抽出する。このように抽出された既決原子は複数あるので、次に S 9 2 7 の処理を行い、抽出された既決原子の中から正準化番号が最小の原子を選択する。即ち、入力番号 2 の原子の正準化番号は 3 で、入力番号 4 の原子の正準化番号は 2 である。このため、入力番号 4 の原子を選択する。そして、S 9 2 8 の処理を行い、入力番号 4 の原子に結合した入力番号 5 の原子に正準化番号 $k = 4$ を与える。

同様の処理を繰り返すことにより、入力番号 1 の原子に正準化番号 1 が、入力番号 6 の原子に正準化番号 6 がそれぞれ与えられる。また、入力番号 7 の原子に正準化番号 7 が、入力番号 8 の原子に正準化番号 8 がそれぞれ与えられる。

その後、S 9 2 3 の処理を行い、この段階では全ての原子の正準化番号が求められているので処理を終了する。その結果、図 3 4 に示すような正準化番号が得られる。

次に、図 3 5 のフローチャートを用いて、S 9 3 0 で呼び出される正準化データ作成ルーチン 9 1 d の処理について説明する。この処理は、まず、図 3 6 A ~ B に示すように入力番号を正準化番号に置き換えて、結合表 8 1 を書き替える (S 9 3 1)。そして、この結合表 8 1 に基づいて、各原子に対して 3 種類のデータ (P_i , T_i , S_i) を求める (S 9 3 2)。ここで、 P_i は正準化番号 i ($i > 1$) の原子に結合して、番号が最小である原子の正準化番号である。また、

T_i は正準化番号 i ($i > 1$) の原子と正準化番号 P_i の原子との結合の種類記号 (この例では単結合は一、二重結合は=、三重結合は#、芳香結合は%など) である。さらに、 S_i は正準化番号 i ($i > 0$) の原子の種類記号 (この例では元素番号) である。

具体的には、まず、正準化番号 1 の原子の元素番号について、原子テーブル 8 1 g を参照して調べる。その結果、 $S_1 = \text{"N"}$ が得られる。次に、正準化番号 2 の原子と結合した原子について、原子対テーブル 8 1 h を参照して調べる。その結果、正準化番号 1, 4 の原子が得られる。これらの原子の中で最小の正準化番号は 1 なので、 $P_2 = 1$ となる。そして、正準化番号 2 の原子と正準化番号 1 の原子との結合は単結合なので、 $T_2 = \text{"-"}$ となる。さらに、原子テーブル 8 1 g を参照することにより、 $S_2 = \text{"C"}$ が得られる。

次に、正準化番号 3 の原子と結合した原子について、原子対テーブル 8 1 h を参照して調べる。その結果、正準化番号 1, 5 の原子が得られる。これらの原子の中で最小の正準化番号は 1 なので、 $P_3 = 1$ となる。そして、正準化番号 3 の原子と正準化番号 1 の原子との結合は二重結合なので、 $T_3 = \text{"="}$ となる。さらに、原子テーブル 8 1 g を参照することにより、 $S_3 = \text{"C"}$ が得られる。以下同様に処理を行うことにより、 $P_4 = 2$ 、 $P_5 = 3$ 、 $P_6 = 4$ 、 $P_7 = 4$ 、 $P_8 = 5$ 、 $T_4 \sim T_8 = \text{"-"}$ 、 $S_4 \sim S_8 = \text{"C"}$ が得られる。

次に、S 9 3 2 の処理で T_i を求めた際に参照されなかった結合原子対を抽出する (S 9 3 3)。この処理は原子対テーブル 8 1 h を参照して行う。その結果、正準化番号 5 の原子と正準化番号 6 の原子との結合原子対が抽出される。そして、抽出された結合原子対に対して 3 種類のデータ (R^1_j , R^2_j , H_j) を求める (S 9 3 4)。ここで、 R^1_j , R^2_j はその結合を構成する 2 つの原子の正準化番号である。また、 H_j はその結合の種類記号 (この例では T_i と同じものを用いる) である。なお、 R^1_j と R^2_j とは、 $R^1_j > R^2_j$ の関係を満たすものとする。また、別の結合原子対 (R^1_k , R^2_k) とは、 $R^1_j \leq R^1_k$ の

関係を満たすか、 $R^1_j = R^1_k$ で且つ $R^2_j < R^2_k$ の関係を満たすものとする。

以上の処理によって、図37に示す正準化木構造データが作成できた。

次に、S932及びS934の処理で求めた各データを一行に並べて、正準化データを作成する(S935)。即ち、原子の種類記号及び結合の種類記号と異なる区切り記号Fを定義して、S932及びS934の処理で求めた各データを以下のように並べる。

$S_1, P_2, T_2, S_2, P_3, T_3, S_3, P_4, T_4, S_4, \dots, P_N, T_N, S_N, F, R^1_1, H_1, R^2_1, F, R^1_2, H_2, R^2_2, \dots, F, R^1_M, H_M, R^2_M, F$

ここで、Nは総原子数であり、MはS934の抽出された結合原子対の総数である。

このようにして得られたデータ列は、化合物の構造と一意に対応する正準化データである。具体的には、区切り記号Fを“/”として、得られたデータを所定の順番に並べると、

“N1=C1=C2-C3-C4-C4-C5-C/5-6/”

が得られる。そして、この正準化データは化合物情報ファイル31に書き込まれて保存される(S936)。その後、処理は終了する。

なお、本発明にかかる正準化データ作成手段及び方法は上記実施形態に限定されることなく、本発明の趣旨から逸脱しない範囲内において、例えば以下のように変更することも可能である。

(1) 上記実施形態では、正準化データとして原子の種類記号 S_i を含めたデータ列を用いているが、最も出現頻度が高い原子の種類記号(通常は炭素のC)をデータ列から除いてもよい。即ち、上述した正準化データから炭素Cの記号を省略することにより、

“N1-1=2-3-4-4-5-/5-6/”

が得られる。このようにしてデータ列を短くすることにより、化合物情報ファイル31に書き込まれるデータ量を削減することができる。

(2) 正準化番号付与ルーチン91cでは、S929の処理でクラス番号 C_i が最大の未決原子が複数選択された場合に、次の処理を追加してもよい。

(a) クラス番号 C_i が最大の未決原子が環状構造部分に属していない場合には、複数の未決原子の中から任意の未決原子を選択して、この未決原子の正準化番号を k とする。その後に処理をS923に戻す。

(b) クラス番号 C_i が最大の未決原子が環状構造部分に属している場合には、S929で選択された未決原子（以下、候補原子という）とこれらの候補原子に結合した既決原子との結合を切断した構造について、各候補原子に対して次のベクトル量を定義する。

m_{ik} : 候補原子 i と、正準化番号 k である原子間の最小結合数

予め、この属性について優先順位を定めておき、最も優先順位の高い原子 i を選択して、その原子の正準化番号を k とする。その後に処理をS923に戻す。

ここで、原子の属性値における優先順位の判定基準は以下の通りである。まず、非ベクトル量については優先順位の大小による。また、ベクトル量については2つのベクトル i , k の要素が属性 V_{ij} , V_{kj} のとき、 $V_{ij} \neq V_{kj}$ である要素の中で最小の j における大小を優先順位の判定基準とする。このような判定基準を用いることにより、属性 b_{ij} , d_{ij} , V_{ij}^n , m_{ij} の優先順位を定めることができる。また複数の属性によって優先順位が定まる場合には、予め、属性間にも優先順位を定めておき、優先順位の高い属性での判定を優先する。

なお、上記の本発明にかかる正準化データ作成方法によって図38Aに示す C_{60} 分子の正準化データを求めたところ、 C_{60} 分子の構造を一意的に特定する正準化データ（図38B）を僅か1.5秒で得ることができた。これに対して、原子を等価原子毎に分類する処理過程を経ないモルガン・アルゴリズムによって同一性能の情報処理装置を用いて C_{60} 分子の正準化データを求めたところ、正準化デ

一タを得るのに550秒を要した。従って、本発明において上記本発明にかかる正準化データ作成手段及び方法を採用すれば、本発明にかかる生化学情報処理の速度が大幅に向上する。

以上、本発明の生化学情報処理装置及び方法の好適な実施態様について説明したが、本発明は上記の実施態様に限定されるものではない。

例えば、本発明にかかる正準化データ作成手段（正準化データ作成プログラム91）は本発明の生化学情報処理装置において他の手段（反応式検出プログラム25等）と一体的に第1記憶装置中に存在しなくてもよく、図39及び図40に示すように本発明にかかる正準化データ作成手段（正準化データ作成プログラム91）と他の手段（反応式検出プログラム25等）とがそれぞれ別個に第1記憶装置20中に存在してもよい。

また、本発明の生化学情報処理装置においては、反応式検出手段（反応式検出プログラム）25、受容体情報検出手段（受容体情報検出プログラム）26及び反応経路検出手段（反応経路検出プログラム）27のうちの全てを備えていなくてもよく、例えば図41に示すように反応式検出手段（反応式検出プログラム）25及び反応経路検出手段（反応経路検出プログラム）27を備えるものであっても、あるいはいずれか一方のみを備えるものであってもよい。この場合、受容体情報ファイル36は必要なく、図2の反応経路図に記載された化合物番号C₁～C₇と酵素番号E₁～E₆との相互の関係は、図42に示す相互関連情報ファイル33に記録されている。詳細には、化合物番号C₁～C₆の各化合物が基質となる酵素の酵素番号E₁～E₆と、化合物番号C₂～C₇の各化合物が生成物となる酵素の酵素番号E₁～E₆と、化合物番号C₆の化合物が阻害する酵素の酵素番号E₄とが、化合物番号C₁～C₇に対応した一覧表として記録されている。このため、化合物番号C₁～C₇をキーとして、相互関連情報ファイル33にアクセスすれば、化合物番号C₁～C₇の各化合物が基質或いは生成物となる酵素の酵素番号E₁～E₆と、化合物番号C₆の化合物が阻害する酵素の酵素番

号E₄とを読み出すことができる。また、この場合のメインプログラム23は、図43に示すように、受容体情報表示プログラムを呼び出すステップS115を含まない以外は図11と同様である。

次に、本発明の実施例に係る生化学情報コンピュータプログラム製品（記録媒体）について説明する。

図44は、本発明の実施例に係る生化学情報コンピュータプログラム製品（記録媒体）2の構成を示すブロック図である。同図に示すように、本実施例の生化学情報記録媒体2は、ファイルを記録するファイル領域2bと、プログラムを記録するプログラム領域2aとを備えている。そして、ファイル領域2bには、化合物情報ファイル31と、酵素情報ファイル32と、相互関連情報ファイル33と、部分相互関連データファイル34と、結合表ファイル35と、受容体情報ファイル36とが記録されている。

この内、化合物情報ファイル31には、化合物番号とこの化合物に対応する正準化データ（カノニカルデータともいう）との関係を示す一覧表、及びこの化合物についての付加情報（参照データともいう）が格納されている。また、酵素情報ファイル32には、酵素番号とこの酵素の基質である化合物の化合物番号とこの酵素による生成物である化合物の化合物番号との関係を示す一覧表、及びこの酵素についての付加情報が格納されている。

さらに、相互関連情報ファイル33には、化合物番号と当該化合物が基質となる酵素の酵素番号と当該化合物が生成物となる酵素の酵素番号と、当該化合物が作動物質となる受容体の番号と当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表が格納されている。さらにまた、部分相互関連データファイル34には反応経路情報が、結合表ファイル35には結合表データがそれぞれ格納できるように用意されている。さらにまた、受容体情報ファイル36には、受容体番号とこの受容体の作動物質である化合物の化合物番号とこの受容体の拮抗物質である化合物の化合物番号との関係を示す一覧表、及びこの受容体についての付

加情報が格納されている。

また、プログラム領域2 aには、生化学情報処理プログラム2 2が記録されている。生化学情報処理プログラム2 2は、処理を統括するメインプログラム2 3と、画像データを3次元表示させる3次元表示プログラム2 4と、化合物間の化学反応式を検出する反応式検出プログラム2 5と、受容体についての付加情報を検出する受容体情報検出プログラム2 6と、複数の化合物の反応経路を検出する反応経路検出プログラム2 7とを備えている。そして、反応式検出プログラム2 5は、第1の処理ルーチン2 5 a～第4の処理ルーチン2 5 dを備え、受容体情報検出プログラム2 6は、第5の処理ルーチン2 6 a～第8の処理ルーチン2 6 dを備え、反応経路検出プログラム2 7は、第9の処理ルーチン2 7 a～第13の処理ルーチン2 7 eを備えている。

生化学情報記録媒体2としては、例えば、フレキシブルディスクやCD-ROMなどの円盤型記録媒体が用いられる。また、磁気テープなどのテープ型記録媒体を用いてもよい。

本実施例の生化学情報記録媒体2は、図4 5及び図4 6に示す情報処理装置1で用いることができる。つまり、情報処理装置1は媒体ドライブ装置3を備えており、この媒体ドライブ装置3に生化学情報記録媒体2を収納することができる。そして、この収納によって、生化学情報記録媒体2に記録された生化学情報は、媒体ドライブ装置3でアクセス可能となる。このため、プログラム領域2 0に記録された生化学情報処理プログラム2 2が情報処理装置1で実行できるようになる。

この情報処理装置1の構成は、次の通りである。まず、上述の媒体ドライブ装置3と、化合物の分子構造図等を示す画像データを記憶する画像メモリ1 0と、オペレーティングシステム(OS)を常駐した作業用メモリ(内部メモリ)1 1と、表示手段であるディスプレイ4 0とを備えている。また、画像データの入力を受け付けるマウス5 1及び記号データの入力を受け付けるキーボード5 2を有

する入力手段である入力装置 50 と、画像データ等を入力するプリンタ 60 と、生化学情報処理プログラム 22 の実行等を制御する CPU 70 とを備えている。

媒体ドライブ装置 3 としては、生化学情報記録媒体 2 に対応して、フレキシブルディスクドライブ装置、CD-ROM ドライブ装置、或いは磁気テープドライブ装置などが用いられる。

なお、本実施例の生化学情報記録媒体 2 に記録されている化合物情報ファイル 31、酵素情報ファイル 32、相互関連情報ファイル 33、部分相互関連データファイル 34、結合表ファイル 35 及び受容体情報ファイル 36 の詳細な構成は前述の通りである（図 2～図 6）。

また、情報処理装置 1 内でのデータの流れも前述の通りであり、入力された画像データ 80 が結合表データ 81、正準化データ 82、3次元データ 83 のいずれにも変換され、用いられるが、プログラム領域 2a に記録された生化学情報処理プログラム 22 では主に正準化データ 82 が用いられていることも前述の通りである（図 7～図 10）。

次に、生化学情報記録媒体 1 のプログラム領域 2a に記録された生化学情報処理プログラム 22 の処理について説明する。この処理は、媒体ドライブ装置 3 により読み出された生化学情報処理プログラム 22 を、実行させることにより行われる。この実行によって、まず、生化学情報処理プログラム 22 のメインプログラム 23 が起動される。

なお、その後のメインプログラム 23、3次元表示プログラム 24、反応式検出プログラム 25、反応経路検出プログラム 27 及び受容体情報検出プログラム 26 による処理の詳細も前述の通りであり（図 11～図 15）、例えば図 16、図 17 に示すように反応式データ等がディスプレイ 40 に表示される。

次に、本発明に好適な正準化データ作成プログラムについて説明する。

図 44 に示す本発明の実施形態である生化学情報コンピュータプログラム製品（記録媒体）2 は本発明にかかる正準化データ作成プログラムを備えており、す

なわち、生化学情報記録媒体2は、ファイルを格納するファイル領域2bと、プログラムを格納するプログラム領域2aとを備えている。そして、ファイル領域2bには、結合表ファイル35、化合物情報ファイル31等が格納されている。

結合表ファイル31には、複数の結合表81を記録することができる。結合表81には化合物を構成する各原子についての固有データ及び原子間の結合対データが記録されており、正準化データ作成プログラム91は結合表81を介してこれらのデータにアクセスできる。

なお、化合物情報ファイル31、結合表81については前述の通りである（図3、図18A及び図18B）。

また、プログラム領域2aには、正準化データ作成プログラム91が格納されている。正準化データ作成プログラム91は、化合物を構成する各原子についての固有データ及び原子間の結合対データに基づいて正準化データを作成するプログラムである。この正準化データ作成プログラム91は、処理を統括するメインルーチン91aと、化合物を構成する各原子にクラス番号を与える構成原子分類ルーチン91bとを備えている。また、正準化データ作成プログラム91は、クラス番号に基づいて各原子に正準化番号を与える正準化番号付与ルーチン91cと、各原子の正準化番号に基づいて正準化データを作成する正準化データ作成ルーチン91dとを備えている。

なお、生化学情報記録媒体2は、前述のように、図45に示す情報処理装置1で利用することができる。なお、ポインティングデバイスには、マウス51以外に、タブレット、ディジタイザ、ライトペンなどがあり、これらのいずれの装置をマウス51の代わりに備えてもよい。

また、情報処理装置1の動作の概要も前述の通りである。

次に、生化学情報記録媒体2のプログラム領域2aに格納された正準化データ作成プログラム91の処理について説明する。この処理は、媒体ドライブ装置3により読み出された正準化データ作成プログラム91を実行させることにより行

われる。この実行によって、まず、正準化データ作成プログラム 91 のメインルーチン 91 a が起動される。

なお、その後のメインルーチン 91 a、構成原子分類ルーチン 91 b、正準化番号付与ルーチン 91 c 及び正準化データ作成ルーチン 91 d による処理の詳細も前述の通りであり（図 20～図 37）、化合物を一意的に特定する正準化データが短時間で得られる。

以上、本発明の生化学情報コンピュータプログラム製品（記録媒体）の好適な実施態様について説明したが、本発明は上記の態様に限定されるものではない。

例えば、本発明にかかる正準化データ作成プログラム 91 は本発明にかかる生化学情報処理プログラム 22 と共に単一の媒体中に存在しなくてもよく、図 47 及び図 48 に示すように本発明にかかる正準化データ作成プログラム 91 と生化学情報処理プログラム 22 とをそれぞれ別個の媒体中に記録してもよい。

すなわち、図 48 に示すように本発明にかかる正準化データ作成プログラム 91 を単独で正準化データ作成用記憶媒体 2 としてもよい。この場合、正準化データ作成用記憶媒体 2 は、図 49 に示す情報処理装置 1 で利用することができる。つまり、情報処理装置 1 は媒体ドライブ装置 3 を備えており、この装置 3 に正準化データ作成用記憶媒体 2 を収納することができる。そして、この収納によって、正準化データ作成用記憶媒体 2 に格納された情報は、媒体ドライブ装置 3 でアクセス可能となる。このため、プログラム領域 2 a に格納された正準化データ作成プログラム 91 が情報処理装置 1 で実行できるようになる。なお、正準化データ作成用記憶媒体 2 としては、例えば、フレキシブルディスクや CD-ROM などの円盤型記憶媒体や、磁気テープなどのテープ型記憶媒体が用いられる。

また、本発明の生化学情報コンピュータプログラム製品（記録媒体）においては、反応式検出プログラム 25、受容体情報検出プログラム 26 及び反応経路検出プログラム 27 のうちの全てを備えていなくてもよく、例えば図 50 に示すように反応式検出手段（反応式検出プログラム）25 及び反応経路検出手段（反応

経路検出プログラム) 27を備えるものであっても、あるいはいずれか一方のみを備えるものであってもよい。この場合、受容体情報ファイル36は必要なく、またこの場合のメインプログラム23は、図43に示すように、受容体情報表示プログラムを呼び出すステップS115を含まない以外は図11と同様である。

なお、本発明は上記実施例に限定されることなく、種々の変形が可能である。例えば、酵素情報ファイル32の参照データの欄に、酵素の構造を規定するアミノ酸配列、あるいは塩基配列を記録してもよい。同様に、受容体情報ファイル36の参照データの欄に、受容体の構造を規定するアミノ酸配列、あるいは塩基配列を記録してもよい。これらの配列を参照データに記録することにより、遺伝情報と関連させて利用することも可能となる。

また、特定の酵素の機能に異常があると、それが代謝異常病として知られる疾患となることがある。そこで、酵素情報ファイル32の参照データの欄に、異常酵素の情報を記録して、代謝異常病の研究に用いてもよい。

さらに、化合物情報ファイル31、酵素情報ファイル32、相互関連情報ファイル33に、生体に異物(本来、生体になく物質)が投与されることにより生じるその異物の変換の情報を記録してもよい。

さらにまた、化合物情報ファイル31、酵素情報ファイル32、相互関連情報ファイル33に、酵素や微生物による物質の生産や変換に関する情報を記録してもよい。

さらにまた、医薬、農薬には、それ自身が酵素阻害剤やアゴニスト(作動物質)、アンタゴニスト(拮抗物質)となっているものも多い。そこで、化合物情報ファイル31に、医薬、農薬の構造や関連する情報を、生体関連物質として記録してもよい。

さらにまた、化合物情報ファイル31の参照データの欄に、化学物質の毒性等、安全性に関する情報を記録して、生体系での物質の挙動と関連させてながら利用してもよい。

さらにまた、化合物情報ファイル31の参照データの欄に、栄養学分野の情報を記録してもよい。

さらにまた、反応経路の表示法については、例えば、全反応経路図をあらかじめ作成し、任意の位置や縮尺で表示し、画面を上下左右にスクロールすることにより、希望する反応経路部分を表示してもよい。化合物の検索では、部分構造による検索（部分一致検索）、あるいは類似性に基づく検索等も採用できる。また、反応経路の検索において、例えばステロイドの代謝のような特定の化合物群を対象としてもよい。

本処理装置または本処理方法は化合物データベースシステムとしても用いることができ、また本発明の媒体に化合物データベースシステムの各情報を記録してもよい。この場合、物性値等の化合物データに基づく検索も可能である。また、化合物の三次元構造データに基づき、分子軌道計算や分子力場計算等の理論化学計算機能を、本処理装置または本処理方法に付加することもできる。本処理装置あるいは本処理方法を用いれば、特定の酵素が阻害、失活された場合、もしくは酵素が欠損している場合の反応経路も知ることができる。

さらにまた、本発明の生化学情報記録媒体には、特定の酵素が阻害、失活された場合、もしくは酵素が欠損している場合の反応経路も知るための情報を記録してもよい。

産業上の利用可能性

以上詳細に説明したように、本発明の生化学情報処理装置及び生化学情報処理方法であれば、反応式の検出、受容体情報の検出及び反応経路の検出を効率よく行うことができる。また、本発明の生化学情報記録媒体を用いることによって、反応式の検出、受容体情報の検出及び反応経路の検出を効率よく行うことができる。

反応式の検出は、まず、化合物情報ファイルに格納された一覧表が参照され、

正準化データに対応する化合物番号が読み出される。次に、この化合物番号に基づいて、相互関連情報ファイルが参照され、この化合物が基質あるいは生成物となる酵素番号が読み出される。さらに、この酵素番号に基づいて、酵素情報ファイルが参照され、この酵素についての情報が読み出される。そして、このように読み出された化合物や酵素の情報から、この化合物が関わる化学反応式が得られる。

このように、化合物情報ファイル、酵素情報ファイル、及び相互関連情報ファイルを相互に参照することにより、化合物の構造をキーとした場合でも、この化合物が基質或いは生成物となる酵素についての各種情報を効率的に獲得することができる。

特に、相互関連情報ファイルには、化合物とこの化合物が基質或いは生成物となる酵素との関係を示す一覧表が格納されているので、基質である化合物と、生成物である化合物と、基質から生成物に変化させる酵素との関係が容易に検索でき、効率的に化学反応式を得ることができる。

また、受容体情報の検出は、まず、化合物情報ファイルに格納された一覧表が参照され、正準化データに対応する化合物番号が読み出される。次に、この化合物番号に基づいて、相互関連情報ファイルが参照され、この化合物が作動物質あるいは拮抗物質となる受容体番号が読み出される。さらに、この受容体番号に基づいて、受容体情報ファイルが参照され、この受容体についての付加情報が読み出される。そして、このように読み出された受容体についての付加情報が表示手段に表示される。

このように、化合物情報ファイル、受容体情報ファイル、及び相互関連情報ファイルを相互に参照することにより、化合物の構造をキーとした場合でも、この化合物が作動物質あるいは拮抗物質となる受容体についての各種情報を効率的に獲得することができる。

特に、相互関連情報ファイルには、化合物とこの化合物が作動物質あるいは拮

抗物質となる受容体との関係を示す一覧表が格納されているので、作動物質である化合物と、拮抗物質である化合物と、受容体との関係が容易に検索でき、効率的に受容体についての各種情報を得ることができる。

さらに、反応経路の検出は、まず、化合物情報ファイルに格納された一覧表が参照され、正準化データに対応する化合物番号が読み出される。次に、この化合物番号に基づいて、相互関連情報ファイルが参照され、この化合物を基質とすると酵素番号とこの化合物を生成物とする酵素番号とがそれぞれ読み出される。さらに、これらの酵素番号に基づいて、酵素情報ファイルが参照され、酵素ごとに基質である化合物番号と生成物である化合物番号とが読み出される。相互関連情報ファイルと酵素情報ファイルとからの呼び出しは、繰り返して行われる。そして、このように読み出された複数の化合物番号及び複数の酵素番号から、これらの化合物の反応経路が得られる。

このように、化合物情報ファイル、酵素情報ファイル、及び相互関連情報ファイルを相互に参照することにより、効率的に複数の化合物が関わる反応経路を検索することができる。

特に、相互関連情報ファイルには、化合物とこの化合物が基質或いは生成物となる酵素との関係を示す一覧表が格納されているので、基質である化合物と、生成物である化合物と、基質から生成物に変化させる酵素との関係が容易に検索でき、効率的に複数の化合物が関わる反応経路を得ることができる。

さらに、本発明にかかる正準化データ作成手段（正準化データ作成プログラム）を採用すれば、入力手段で受け付けられた各原子についての固有データ及び原子間の結合対データは正準化データ作成手段に与えられる。そして、正準化データ作成手段では、これらのデータに基づいて正準化データが短時間に作成される。また、本発明にかかる正準化データ作成方法によれば、化合物を構成する各原子についての固有データ及び原子間の結合対データに基づいて正準化データが短時間に作成される。このように、本発明にかかる正準化データ作成手段（正準化デ

ータ作成プログラム) 及び正準化データ作成方法で作成された正準化データは、非常に短い文字・数字・記号列であり、少ない記憶領域で正準化データを保存することができる。そのため、本発明にかかる正準化データ作成手段(正準化データ作成プログラム) 及び正準化データ作成方法を化合物/反応データベースシステムで利用すれば、化合物/反応データベースシステムの記憶領域の使用量を大幅に削減させることができる。

請 求 の 範 囲

1. 化合物と酵素とについての生化学情報が記憶された記憶手段と、

前記生化学情報を表す画像データまたは前記生化学情報を表す記号データの入力を受け付ける入力手段と、

基質および／または生成物である化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物が関わる化学反応式を検出する反応式検出手段と、

少なくとも化学反応式の反応式図を表示する表示手段とを備える生化学情報処理装置であって；

前記記憶手段は、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及び当該化合物についての付加情報が格納された化合物情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物である化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納された酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化合物が生成物となる酵素の番号との関係を示す一覧表が格納された相互関連情報ファイルとを備え；

前記反応式検出手段は、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらに正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第1の処理部と、

前記第1の処理部で読み出された化合物の番号に基づいて、化合物が基質ある

いは生成物となる酵素の番号を前記相互関連情報ファイルから読み出す第2の処理部と、

前記第2の処理部で読み出された番号の酵素と前記化合物と共に反応系を構成する他の化合物の番号、及び前記酵素についての付加情報を前記酵素情報ファイルから読み出す第3の処理部と、

前記第1の処理部で読み出された化合物の番号と、前記第2の処理部で読み出された酵素の番号と、前記第3の処理部で読み出された他の化合物の番号とから、前記入力手段で受け付けられた化合物の反応式図を前記表示手段に表示し、さらに、前記第3の処理部で読み出された酵素についての付加情報を前記表示手段に表示する第4の処理部とを備える；

生化学情報処理装置。

2. 前記生化学情報処理装置が、化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物が作動物質および／または拮抗物質となる受容体についての付加情報を検出する受容体情報検出手段を更に備えており；

前記記憶手段が、受容体についての生化学情報を更に記憶しており、かつ

受容体の番号と当該受容体の作動物質および／または拮抗物質である化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された受容体情報ファイルを更に備えており；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表を格納しており；

前記受容体情報検出手段は、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づ

いて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第5の処理部と、

前記第5の処理部で読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す第6の処理部と、

前記第6の処理部で読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す第7の処理部と、

前記第7の処理部で読み出された受容体についての少なくとも付加情報を前記表示手段に表示する第8の処理部とを備える；

請求項1に記載の生化学情報処理装置。

3. 前記生化学情報処理装置が、反応経路を構成する複数の化合物の中から選択される所定の化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記複数の化合物の反応経路を検出する反応経路検出手段を更に備えており；

前記反応経路検出手段は、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第9の処理部と、

前記第9の処理部で読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す第10の処理部と、

前記第10の処理部で読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報フ

ファイルから読み出す第 11 の処理部と、

前記第 10 の処理部による処理と前記第 11 の処理部による処理とを繰り返し、
所定反応経路内の化合物及び酵素を検索する第 12 の処理部と、

前記第 10 の処理部で読み出された酵素の番号と、前記第 11 の処理部で読み
出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前
記表示手段に表示する第 13 の処理部とを備える；

請求項 1 又は 2 に記載の生化学情報処理装置。

4. 化合物と酵素とについての生化学情報が記憶された記憶手段と、

前記生化学情報を表す画像データまたは前記生化学情報を表す記号データの入
力を受け付ける入力手段と、

反応経路を構成する複数の化合物の中から選択される所定の化合物についての
データが前記入力手段で受け付けられた場合に、このデータに基づいて前記複数
の化合物の反応経路を検出する反応経路検出手段と、

少なくとも化学反応式の反応式図を表示する表示手段とを備える生化学情報処
理装置であって；

前記記憶手段は、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及
び当該化合物についての付加情報が格納された化合物情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物であ
る化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納
された酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化
合物が生成物となる酵素の番号との関係を示す一覧表が格納された相互関連情報
ファイルとを備え；

前記反応経路検出手段は、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学

構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第9の処理部と、

前記第9の処理部で読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す第10の処理部と、

前記第10の処理部で読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報ファイルから読み出す第11の処理部と、

前記第10の処理部による処理と前記第11の処理部による処理とを繰り返し、所定反応経路内の化合物及び酵素を検索する第12の処理部と、

前記第10の処理部で読み出された酵素の番号と、前記第11の処理部で読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前記表示手段に表示する第13の処理部とを備える；

生化学情報処理装置。

5. 前記生化学情報処理装置が、化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物が作動物質および／または拮抗物質となる受容体についての付加情報を検出する受容体情報検出手段を更に備えており；

前記記憶手段が、受容体についての生化学情報を更に記憶しており、かつ

受容体の番号と当該受容体の作動物質および／または拮抗物質である化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された受容体情報ファイルを更に備えており；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動

物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表を格納しており；

前記受容体情報検出手段は、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第5の処理部と、

前記第5の処理部で読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す第6の処理部と、

前記第6の処理部で読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す第7の処理部と、

前記第7の処理部で読み出された受容体についての少なくとも付加情報を前記表示手段に表示する第8の処理部とを備える；

請求項4に記載の生化学情報処理装置。

6. 前記入力手段が、化合物を構成する各原子についての固有データ及び原子間の結合対データの入力を受け付け；

前記生化学情報処理装置が、前記入力手段で受け付けられた各データに基づいて、前記化合物の化学構造を一意的に特定できる正準化データを作成する正準化データ作成手段を更に備えており；

前記正準化データ作成手段は、

前記入力手段で受け付けられた各データに基づいて、各原子を等価原子ごとに別のクラスに分類して、クラスごとに異なるクラス番号を各原子に与える構成原子分類処理部と、

前記構成原子分類処理部で各原子に与えられたクラス番号に基づいて、前記化

化合物の構造と一意に対応した正準化番号を各原子に与える正準化番号付与処理部と、

前記正準化番号付与処理部で各原子に与えられた正準化番号に基づいて、前記正準化データを作成する正準化データ作成処理部とを備える；

請求項1～5のうちのいずれかに記載の生化学情報処理装置。

7. 前記構成原子分類処理部は、

各原子に3種類の属性 (a_i , b_{ij} , d_{ij}) を与え、これらの属性が一つでも異なる原子は非等価であると判定できることを利用して、各原子を等価原子毎に異なるクラス番号を付与しており、

前記3種類の属性 (a_i , b_{ij} , d_{ij}) の中で、 a_i は入力番号 i の原子の種類番号であり、 b_{ij} は入力番号 i の原子に隣接する結合のうち、その種類番号が j である結合の数であり、 d_{ij} は入力番号 i の原子から最短経路により j 個の結合を経て巡れる道筋の数であり；

前記正準化番号付与処理部は、

正準化番号を1から昇順に各原子に与える過程において、前記クラス番号の優先順位が最高である原子に正準化番号1を与え、以降正準化番号 n までが付与されている時、既に正準化番号が与えられている原子でかつ未だ正準化番号が与えられていない原子が結合している原子の中で正準化番号が最小である原子を選び、その原子に結合している未だ正準化番号が与えられていない原子の中で前記クラス番号の優先順位が最高である原子に正準化番号 $n+1$ を与えており；

前記正準化データ作成処理部は、

各原子に3種類の属性 (P_i , T_i , S_i) を与えて、これらの属性を一行に並べることによって前記正準化データを作成しており、

前記3種類の属性 (P_i , T_i , S_i) の中で、 P_i は正準化番号 i の原子に結合し且つ正準化番号が最小の原子の正準化番号であり、 T_i は正準化番号 i の原子と正準化番号 P_i の原子との結合の種類記号であり、 S_i は正準化番号 i の

原子の種類記号である；

請求項 6 に記載の生化学情報処理装置。

8. 化合物と酵素とについての生化学情報が記憶された記憶手段と、
前記生化学情報を表す画像データまたは前記生化学情報を表す記号データの入
力を受け付ける入力手段と、

少なくとも化学反応式の反応式図を表示する表示手段とを備える情報処理装置
を用いた生化学情報処理方法であって；

前記記憶手段は、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及
び当該化合物についての付加情報が格納された化合物情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物であ
る化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納
された酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化
合物が生成物となる酵素の番号との関係を示す一覧表が格納された相互関連情報
ファイルとを備え；

前記生化学情報処理方法は、

基質および／または生成物である化合物についてのデータが前記入力手段で受
け付けられた場合に、このデータから当該化合物の化学構造を一意的に示す前記
正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファ
イルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、
前記正準化データに対応する化合物の番号を読み出す第 1 のステップと、

前記第 1 のステップで読み出された化合物の番号に基づいて、この化合物が基
質あるいは生成物となる酵素の番号を前記相互関連情報ファイルから読み出す第
2 のステップと、

前記第 2 のステップで読み出された番号の酵素と前記化合物と共に反応系を構

成する他の化合物の番号、及び前記酵素についての付加情報を、前記酵素情報ファイルから読み出す第3のステップと、

前記第1のステップで読み出された化合物の番号と、前記第2のステップで読み出された酵素の番号と、前記第3のステップで読み出された他の化合物の番号とから、前記入力手段で受け付けられた化合物の反応式図を前記表示手段に表示し、さらに、前記第3のステップで読み出された酵素についての付加情報を前記表示手段に表示する第4のステップとを備える；

生化学情報処理方法。

9. 前記記憶手段が、受容体についての生化学情報を更に記憶しており、かつ

受容体の番号と当該受容体の作動物質および／または拮抗物質である化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された受容体情報ファイルを更に備えており；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表を格納しており；

前記生化学情報処理方法が、

化合物についてのデータが前記入力手段で受け付けられた場合に、このデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第5のステップと、

前記第5のステップで読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す第6のステップと、

前記第 6 のステップで読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す第 7 のステップと、

前記第 7 のステップで読み出された受容体についての少なくとも付加情報を前記表示手段に表示する第 8 のステップとを更に備える；

請求項 8 に記載の生化学情報処理方法。

10. 前記生化学情報処理方法が、

反応経路を構成する複数の化合物の中から選択される所定の化合物についてのデータが前記入力手段で受け付けられた場合に、このデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第 9 のステップと、

前記第 9 のステップで読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す第 10 のステップと、

前記第 10 のステップで読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報ファイルから読み出す第 11 のステップと、

前記第 10 のステップと前記第 11 のステップとを繰り返し、所定反応経路内の化合物及び酵素を検索する第 12 のステップと、

前記第 10 のステップで読み出された酵素の番号と、前記第 11 のステップで読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前記表示手段に表示する第 13 のステップとを更に備える；

請求項 8 又は 9 に記載の生化学情報処理方法。

11. 化合物と酵素とについての生化学情報が記憶された記憶手段と、

前記生化学情報を表す画像データまたは前記生化学情報を表す記号データの入

力を受け付ける入力手段と、

少なくとも化学反応式の反応式図を表示する表示手段とを備える情報処理装置を用いた生化学情報処理方法であって；

前記記憶手段は、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及び当該化合物についての付加情報が格納された化合物情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物である化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納された酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化合物が生成物となる酵素の番号との関係を示す一覧表が格納された相互関連情報ファイルとを備え；

前記生化学情報処理方法は、

反応経路を構成する複数の化合物の中から選択される所定の化合物についてのデータが前記入力手段で受け付けられた場合に、このデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第9のステップと、

前記第9のステップで読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す第10のステップと、

前記第10のステップで読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報ファイルから読み出す第11のステップと、

前記第10のステップと前記第11のステップとを繰り返し、所定反応経路内

の化合物及び酵素を検索する第12のステップと、

前記第10のステップで読み出された酵素の番号と、前記第11のステップで読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前記表示手段に表示する第13のステップとを備える；

生化学情報処理方法。

12. 前記記憶手段が、受容体についての生化学情報を更に記憶しており、かつ

受容体の番号と当該受容体の作動物質および／または拮抗物質である化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された受容体情報ファイルを更に備えており；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表を格納しており；

前記生化学情報処理方法が、

化合物についてのデータが前記入力手段で受け付けられた場合に、このデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す第5のステップと、

前記第5のステップで読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す第6のステップと、

前記第6のステップで読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す第7のステップと、

前記第7のステップで読み出された受容体についての少なくとも付加情報を前

記表示手段に表示する第 8 のステップとを更に備える；

請求項 1 1 に記載の生化学情報処理方法。

1 3. 前記入力手段が、化合物を構成する各原子についての固有データ及び原子間の結合対データの入力を受け付け；

前記生化学情報処理方法が、

前記入力手段で受け付けられた各データに基づいて、各原子を等価原子ごとに別のクラスに分類して、クラスごとに異なるクラス番号を各原子に与える構成原子分類ステップと、

前記構成原子分類ステップで各原子に与えられたクラス番号に基づいて、前記化合物の構造と一意的に対応した正準化番号を各原子に与える正準化番号付与ステップと、

前記正準化番号付与ステップで各原子に与えられた正準化番号に基づいて、前記化合物の化学構造を一意的に特定できる正準化データを作成する正準化データ作成ステップとを更に備える；

請求項 8 ～ 1 2 のうちのいずれかに記載の生化学情報処理方法。

1 4. 前記構成原子分類ステップは、

各原子に 3 種類の属性 (a_i , b_{ij} , d_{ij}) を与え、これらの属性が一つでも異なる原子は非等価であると判定できることを利用して、各原子を等価原子毎に異なるクラス番号を付与しており、

前記 3 種類の属性 (a_i , b_{ij} , d_{ij}) の中で、 a_i は入力番号 i の原子の種類番号であり、 b_{ij} は入力番号 i の原子に隣接する結合のうち、その種類番号が j である結合の数であり、 d_{ij} は入力番号 i の原子から最短経路により j 個の結合を経て巡れる道筋の数であり；

前記正準化番号付与ステップは、

正準化番号を 1 から昇順に各原子に与える過程において、前記クラス番号の優先順位が最高である原子に正準化番号 1 を与え、以降正準化番号 n までが付与さ

れている時、既に正準化番号が与えられている原子でかつ未だ正準化番号が与えられていない原子が結合している原子の中で正準化番号が最小である原子を選び、その原子に結合している未だ正準化番号が与えられていない原子の中で前記クラス番号の優先順位が最高である原子に正準化番号 $n+1$ を与えており；

前記正準化データ作成ステップは、

各原子に3種類の属性 (P_i , T_i , S_i) を与えて、これらの属性を一列に並べることによって前記正準化データを作成しており、

前記3種類の属性 (P_i , T_i , S_i) の中で、 P_i は正準化番号 i の原子に結合し且つ正準化番号が最小の原子の正準化番号であり、 T_i は正準化番号 i の原子と正準化番号 P_i の原子との結合の種類記号であり、 S_i は正準化番号 i の原子の種類記号である；

請求項13に記載の生化学情報処理方法。

15. 生化学情報を表す画像データまたは生化学情報を表す記号データの入力を受け付ける入力手段と、少なくとも化学反応式の反応式図を表示する表示手段と、コンピュータ使用可能な媒体から情報を読み出す読出手段とを備える情報処理装置と共に用いるための生化学情報コンピュータプログラム製品であって；

ファイルを記録するためのファイル領域とプログラムを記録するためのプログラム領域とを有しており、かつ、前記入力手段から入力されたデータに基づいて少なくとも反応式図を効率的に検索させかつ前記表示手段で表示させるための該媒体中に具現化されたコンピュータ読取可能なファイル及びプログラムを有しているコンピュータ使用可能な媒体を前記コンピュータプログラム製品は備えており；

該コンピュータプログラム製品は、

前記ファイル領域において、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及び当該化合物についての付加情報が格納された、コンピュータ読取可能な化合物

情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物である化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納された、コンピュータ読取可能な酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化合物が生成物となる酵素の番号との関係を示す一覧表が格納された、コンピュータ読取可能な相互関連情報ファイルとを有しており、

前記プログラム領域において、

基質および／または生成物である化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物に関わる化学反応式を検出する、コンピュータ読取可能な反応式検出プログラムを有しており；

前記反応式検出プログラムは、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらに正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す、コンピュータ読取可能な第1の処理ルーチンと、

前記第1の処理ルーチンで読み出された化合物の番号に基づいて、化合物が基質あるいは生成物となる酵素の番号を前記相互関連情報ファイルから読み出す、コンピュータ読取可能な第2の処理ルーチンと、

前記第2の処理ルーチンで読み出された番号の酵素と前記化合物と共に反応系を構成する他の化合物の番号、及び前記酵素についての付加情報を前記酵素情報ファイルから読み出す、コンピュータ読取可能な第3の処理ルーチンと、

前記第1の処理ルーチンで読み出された化合物の番号と、前記第2の処理ルーチンで読み出された酵素の番号と、前記第3の処理ルーチンで読み出された他の化合物の番号とから、前記入力手段で受け付けられた化合物の反応式図を前記表

示手段に表示し、さらに、前記第3の処理ルーチンで読み出された酵素についての付加情報を前記表示手段に表示する、コンピュータ読取可能な第4の処理ルーチンとを備える；

生化学情報コンピュータプログラム製品。

16. 前記コンピュータプログラム製品が、前記ファイル領域において、受容体の番号と当該受容体の作動物質および／または拮抗物質である化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された、コンピュータ読取可能な受容体情報ファイルを更に有しており；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表を格納しており；

前記コンピュータプログラム製品が、前記プログラム領域において、化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物が作動物質および／または拮抗物質となる受容体についての付加情報を検出する、コンピュータ読取可能な受容体情報検出プログラムを更に有しており；

前記受容体情報検出プログラムは、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す、コンピュータ読取可能な第5の処理ルーチンと、

前記第5の処理ルーチンで読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す、コンピュータ読取可能な第6の処理ルーチンと、

前記第 6 の処理ルーチンで読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す、コンピュータ読取可能な第 7 の処理ルーチンと、

前記第 7 の処理ルーチンで読み出された受容体についての少なくとも付加情報を前記表示手段に表示する、コンピュータ読取可能な第 8 の処理ルーチンとを備える；

請求項 15 に記載の生化学情報コンピュータプログラム製品。

17. 前記コンピュータプログラム製品が、前記プログラム領域において、
反応経路を構成する複数の化合物の中から選択される所定の化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記複数の化合物の反応経路を検出する、コンピュータ読取可能な反応経路検出プログラムを更に有しており；

前記反応経路検出プログラムは、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す、コンピュータ読取可能な第 9 の処理ルーチンと、

前記第 9 の処理ルーチンで読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す、コンピュータ読取可能な第 10 の処理ルーチンと、

前記第 10 の処理ルーチンで読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報ファイルから読み出す、コンピュータ読取可能な第 11 の処理ルーチンと、

前記第 10 の処理ルーチンによる処理と前記第 11 の処理ルーチンによる処理とを繰り返し、所定反応経路内の化合物及び酵素を検索する、コンピュータ読取

可能な第 12 の処理ルーチンと、

前記第 10 の処理ルーチンで読み出された酵素の番号と、前記第 11 の処理ルーチンで読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前記表示手段に表示する、コンピュータ読取可能な第 13 の処理ルーチンとを備える；

請求項 15 又は 16 に記載の生化学情報コンピュータプログラム製品。

18. 生化学情報を表す画像データまたは生化学情報を表す記号データの入力を受け付ける入力手段と、少なくとも化学反応式の反応式図を表示する表示手段と、コンピュータ使用可能な媒体から情報を読み出す読出手段とを備える情報処理装置と共に用いるための生化学情報コンピュータプログラム製品であって；

ファイルを記録するためのファイル領域とプログラムを記録するためのプログラム領域とを有しており、かつ、前記入力手段から入力されたデータに基づいて少なくとも反応式図を効率的に検索させかつ前記表示手段で表示させるための該媒体中に具現化されたコンピュータ読取可能なファイル及びプログラムを有しているコンピュータ使用可能な媒体を前記コンピュータプログラム製品は備えており；

該コンピュータプログラム製品は、

前記ファイル領域において、

化合物の番号と当該化合物に対応する正準化データとの関係を示す一覧表、及び当該化合物についての付加情報が格納された、コンピュータ読取可能な化合物情報ファイルと、

酵素の番号と当該酵素の基質である化合物の番号と当該酵素による生成物である化合物の番号との関係を示す一覧表、及び当該酵素についての付加情報が格納された、コンピュータ読取可能な酵素情報ファイルと、

キーとしての化合物の番号と、当該化合物が基質となる酵素の番号及び当該化合物が生成物となる酵素の番号との関係を示す一覧表が格納された、コンピュ

タ読取可能な相互関連情報ファイルとを有しており、

前記プログラム領域において、

反応経路を構成する複数の化合物の中から選択される所定の化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記複数の化合物の反応経路を検出する、コンピュータ読取可能な反応経路検出プログラムを有しており；

前記反応経路検出プログラムは、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す、コンピュータ読取可能な第 9 の処理ルーチンと、

前記第 9 の処理ルーチンで読み出された化合物の番号に基づいて、この化合物を基質とする酵素の番号とこの化合物を生成物とする酵素の番号とを前記相互関連情報ファイルから読み出す、コンピュータ読取可能な第 10 の処理ルーチンと、

前記第 10 の処理ルーチンで読み出された酵素の番号に基づいて、この酵素の基質である化合物の番号とこの酵素による生成物である化合物の番号を前記酵素情報ファイルから読み出す、コンピュータ読取可能な第 11 の処理ルーチンと、

前記第 10 の処理ルーチンによる処理と前記第 11 の処理ルーチンによる処理とを繰り返し、所定反応経路内の化合物及び酵素を検索する、コンピュータ読取可能な第 12 の処理ルーチンと、

前記第 10 の処理ルーチンで読み出された酵素の番号と、前記第 11 の処理ルーチンで読み出された化合物の番号とから、これらの化合物の反応式図を反応経路に沿って前記表示手段に表示する、コンピュータ読取可能な第 13 の処理ルーチンとを備える；

生化学情報コンピュータプログラム製品。

19. 前記コンピュータプログラム製品が、前記ファイル領域において、受容体の番号と当該受容体の作動物質および／または拮抗物質である化合物の番号との関係を示す一覧表、及び当該受容体についての付加情報が格納された、コンピュータ読取可能な受容体情報ファイルを更に有しており；

前記相互関連情報ファイルが、キーとしての化合物の番号と、当該化合物が基質となる酵素の番号、当該化合物が生成物となる酵素の番号、当該化合物が作動物質となる受容体の番号及び当該化合物が拮抗物質となる受容体の番号との関係を示す一覧表を格納しており；

前記コンピュータプログラム製品が、前記プログラム領域において、化合物についてのデータが前記入力手段で受け付けられた場合に、このデータに基づいて前記化合物が作動物質および／または拮抗物質となる受容体についての付加情報を検出する、コンピュータ読取可能な受容体情報検出プログラムを更に有しており；

前記受容体情報検出プログラムは、

前記入力手段で受け付けられた化合物についてのデータから当該化合物の化学構造を一意的に示す前記正準化データを作成し、さらにこの正準化データに基づいて前記化合物情報ファイルを検索して、前記正準化データが前記化合物情報ファイルに存在する場合に、前記正準化データに対応する化合物の番号を読み出す、コンピュータ読取可能な第5の処理ルーチンと、

前記第5の処理ルーチンで読み出された化合物の番号に基づいて、化合物が作動物質あるいは拮抗物質となる受容体の番号を前記相互関連情報ファイルから読み出す、コンピュータ読取可能な第6の処理ルーチンと、

前記第6の処理ルーチンで読み出された番号の受容体についての少なくとも付加情報を前記受容体情報ファイルから読み出す、コンピュータ読取可能な第7の処理ルーチンと、

前記第7の処理ルーチンで読み出された受容体についての少なくとも付加情報

を前記表示手段に表示する、コンピュータ読取可能な第8の処理ルーチンとを備える；

請求項18に記載の生化学情報コンピュータプログラム製品。

20. 前記入力手段が、化合物を構成する各原子についての固有データ及び原子間の結合対データの入力を受け付け；

前記コンピュータプログラム製品が、前記プログラム領域において、

前記入力手段で受け付けられた各データに基づいて、前記化合物の化学構造を一意的に特定できる正準化データを作成する、コンピュータ読取可能な正準化データ作成プログラムを更に有しており；

前記正準化データ作成プログラムは、

各原子を等価原子ごとに別のクラスに分類して、クラスごとに異なるクラス番号を各原子に与える、コンピュータ読取可能な構成原子分類ルーチンと、

前記構成原子分類ルーチンで各原子に与えられたクラス番号に基づいて、前記化合物の構造と一意的に対応した正準化番号を各原子に与える、コンピュータ読取可能な正準化番号付与ルーチンと、

前記正準化番号付与ルーチンで各原子に与えられた正準化番号に基づいて、前記正準化データを作成する、コンピュータ読取可能な正準化データ作成ルーチンとを備える；

請求項15～19のうちのいずれかに記載の生化学情報コンピュータプログラム製品。

21. 前記構成原子分類ルーチンは、

各原子に3種類の属性 (a_i , b_{ij} , d_{ij}) を与え、これらの属性が一つでも異なる原子は非等価であると判定できることを利用して、各原子を等価原子毎に異なるクラス番号を付与しており、

前記3種類の属性 (a_i , b_{ij} , d_{ij}) の中で、 a_i は入力番号 i の原子の種類番号であり、 b_{ij} は入力番号 i の原子に隣接する結合のうち、その種類番号が

j である結合の数であり、 d_{ij} は入力番号 i の原子から最短経路により j 個の結合を経て巡れる道筋の数であり；

前記正準化番号付与ルーチンは、

正準化番号を 1 から昇順に各原子に与える過程において、前記クラス番号の優先順位が最高である原子に正準化番号 1 を与え、以降正準化番号 n までが付与されている時、既に正準化番号が与えられている原子でかつ未だ正準化番号が与えられていない原子が結合している原子の中で正準化番号が最小である原子を選び、その原子に結合している未だ正準化番号が与えられていない原子の中で前記クラス番号の優先順位が最高である原子に正準化番号 $n+1$ を与えており；

前記正準化データ作成ルーチンは、

各原子に 3 種類の属性 (P_i , T_i , S_i) を与えて、これらの属性を一行に並べることによって前記正準化データを作成しており、

前記 3 種類の属性 (P_i , T_i , S_i) の中で、 P_i は正準化番号 i の原子に結合し且つ正準化番号が最小の原子の正準化番号であり、 T_i は正準化番号 i の原子と正準化番号 P_i の原子との結合の種類記号であり、 S_i は正準化番号 i の原子の種類記号である；

請求項 20 に記載の生化学情報コンピュータプログラム製品。

22. 前記コンピュータ使用可能な媒体が、円盤型記録媒体、または、テープ型記録媒体である、請求項 15～21 のうちのいずれかに記載の生化学情報コンピュータプログラム製品。

図 1

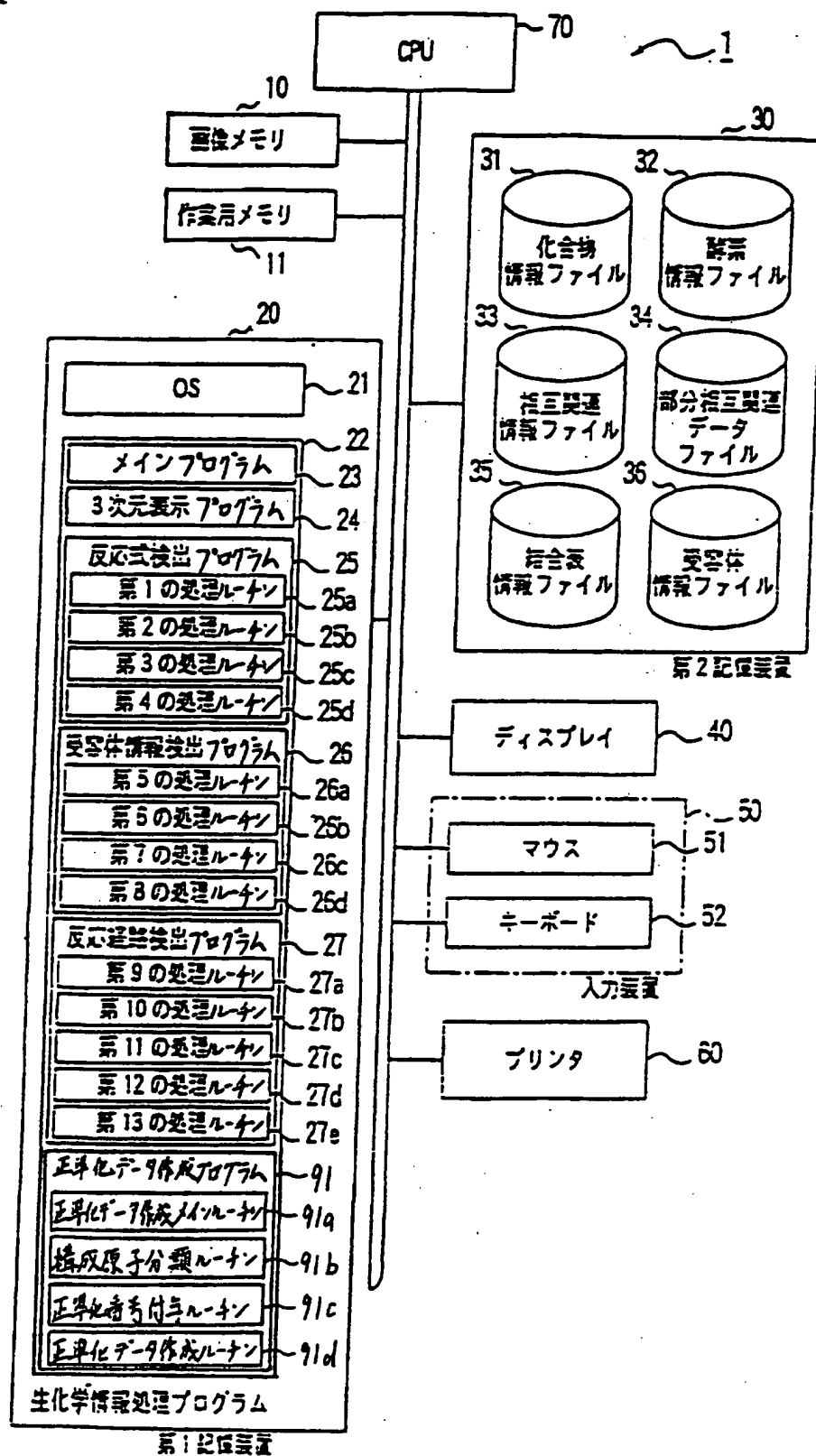


図2

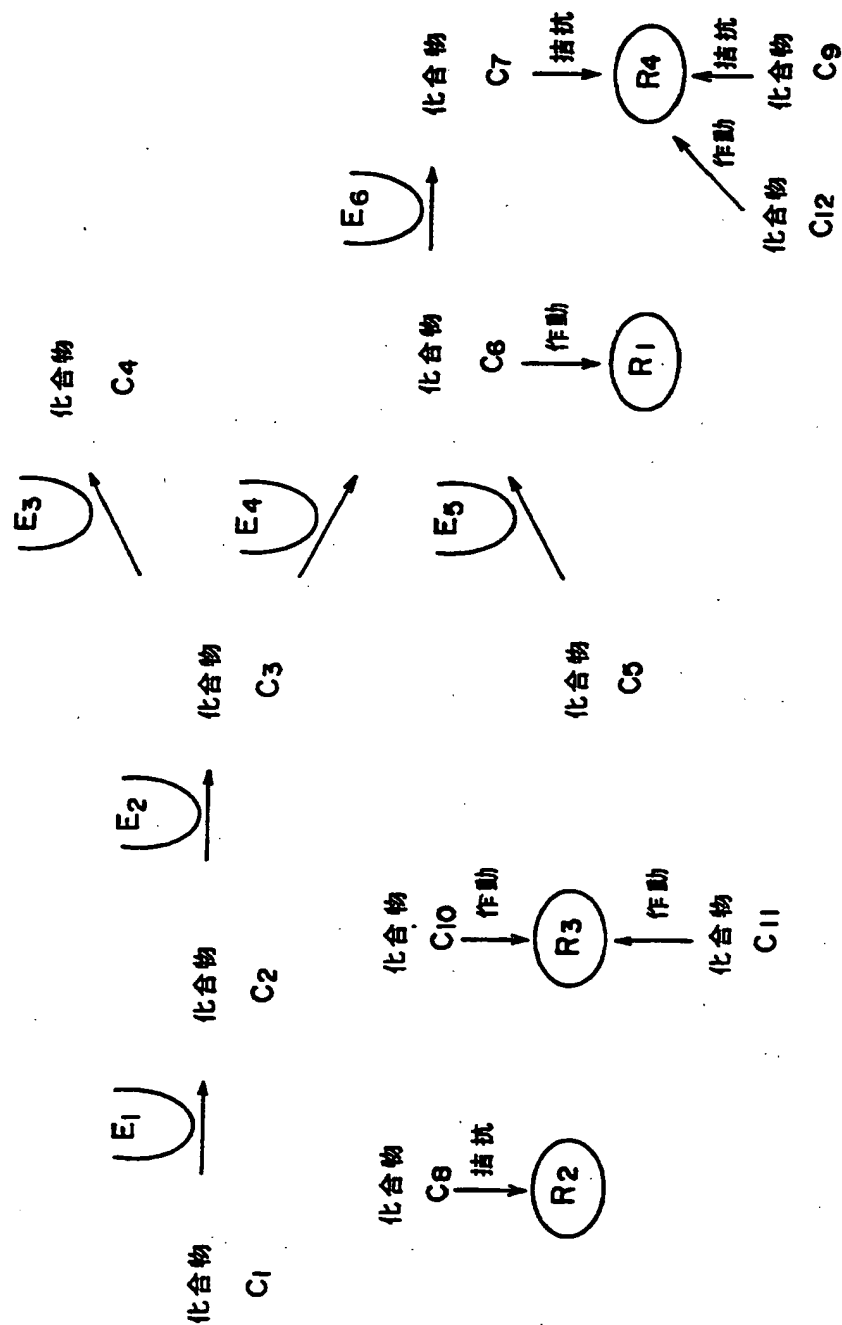


図 3

化合物情報ファイル

31
S

化合物番号	正準化データ	参照データ
C ₁	1%2%	名称、文献、物性等
C ₂	
C ₃	
C ₄	
C ₅	
C ₆	
C ₇	

図 4

酵素情報ファイル

32
S

酵素番号	(基質) 化合物番号	(生成物) 化合物番号	参照データ
E ₁	C ₁	C ₂	名称、文献、物性等
E ₂	C ₂	C ₃	
E ₃	C ₃	C ₄	
E ₄	C ₃	C ₆	
E ₅	C ₅	C ₆	
E ₆	C ₆	C ₇	

図 5

36

受容体情報ファイル

受容体番号	(作動物質) 化合物番号	(拮抗物質) 化合物番号	参照データ
R1	C6		名称、文献、物性、 作用等
R2		C8	
R3	C10,C11		
R4	C12	C7,C9	

図 6

33

相互関連情報ファイル

化合物番号	(基質) 酵素番号	(生成物) 酵素番号	(阻害) 酵素番号	(作動) 受容体番号	(拮抗) 受容体番号
C1	E1				
C2	E2	E1			
C3	E3,E4	E2			
C4		E3			
C5	E5				
C6	E6	E5	E4	R1	
C7		E6			R4
C8					R2
C9					R4
C10				R3	
C11				R3	
C12				R4	

図 7

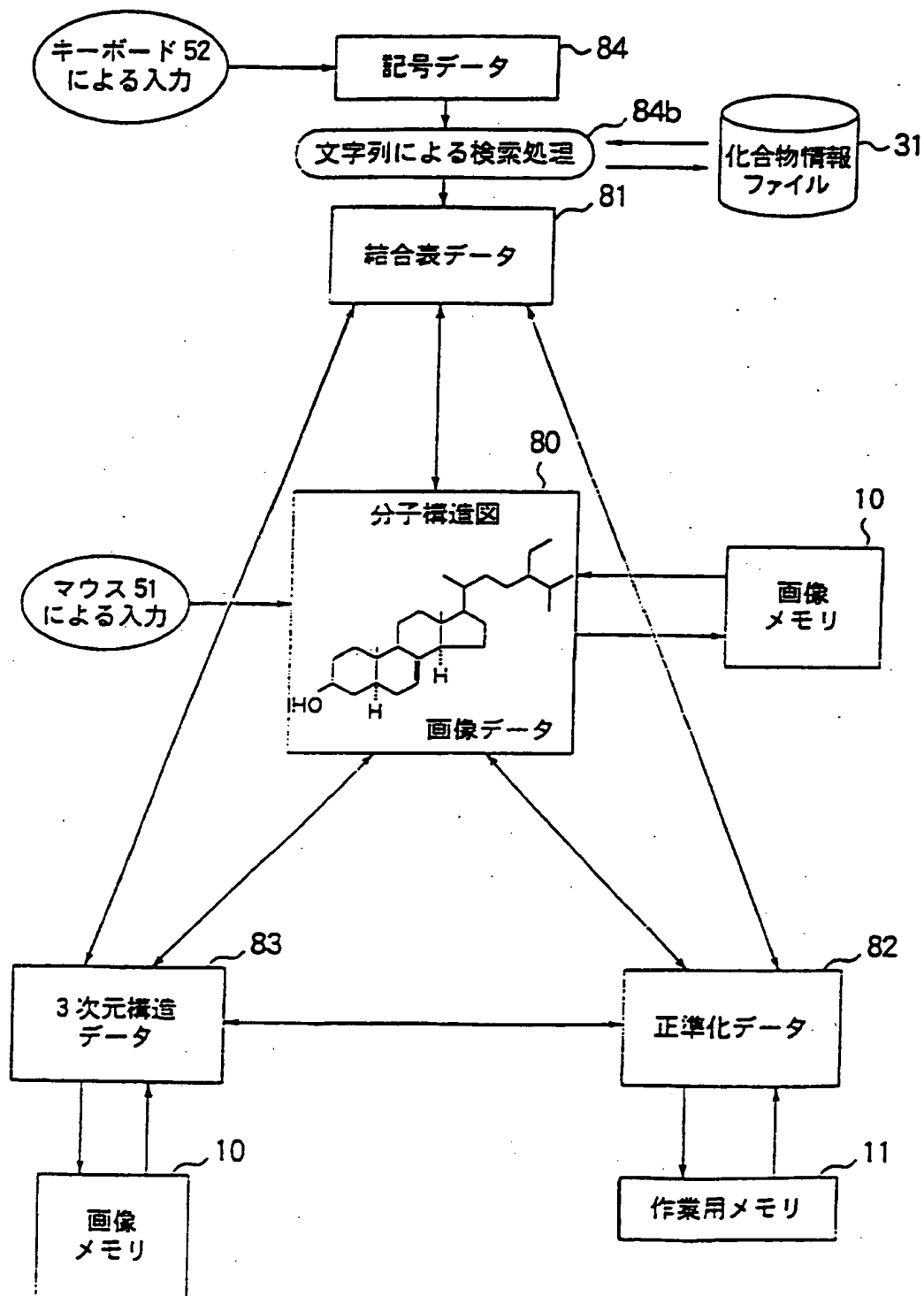


図 8 A 画像データ 80a (分子構造図)

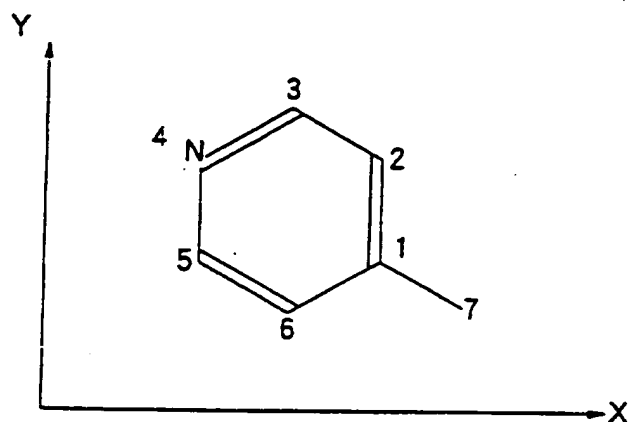


図 8 B 結合表データ 81a

原子数	結合数								
7	7								
原子の番号	X座標	Y座標	Z座標	元素名	電荷	質量	結合原子対	結合種	UP/DOWN
1	3.0386	-2.4082	0.0000	C	0	0	1 2	2	0
2	3.0386	-0.9082	0.0000	C	0	0	2 3	1	0
3	1.7396	-0.1582	0.0000	C	0	0	3 4	2	0
4	0.4405	-0.9082	0.0000	N	0	0	4 5	1	0
5	0.4405	-2.4082	0.0000	C	0	0	5 6	2	0
6	1.7396	-3.1582	0.0000	C	0	0	6 1	1	0
7	4.3375	-3.1582	0.0000	C	0	0	1 7	1	0

図 8 C 正準化データ 82a

1%1%1-2%3%5%N/6%7/

図 9 A 画像データ 80b (分子構造図)

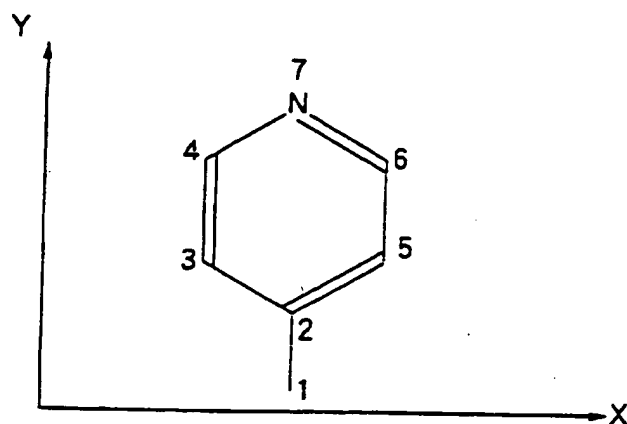


図 9 B 結合表データ 81b

原子数	結合数								
7	7								
原子の番号	X座標	Y座標	Z座標	元素名	電荷	質量	結合原子対	結合種	UP/DOWN
1	3.5200	-4.7352	0.0000	C	0	0	1 2	1	0
2	3.5514	-3.2366	0.0000	C	0	0	2 3	1	0
3	4.8659	-2.5136	0.0000	C	0	0	3 4	2	0
4	4.8973	-1.0143	0.0000	C	0	0	2 5	2	0
5	2.2684	-2.4595	0.0000	C	0	0	5 6	1	0
6	2.2998	-0.9599	0.0000	C	0	0	4 7	1	0
7	3.6143	-0.2372	0.0000	N	0	0	7 6	2	0

図 9 C 正準化データ 82b

1%1%1-2%3%5%N/6%7/

図 1 1

メインプログラム 23

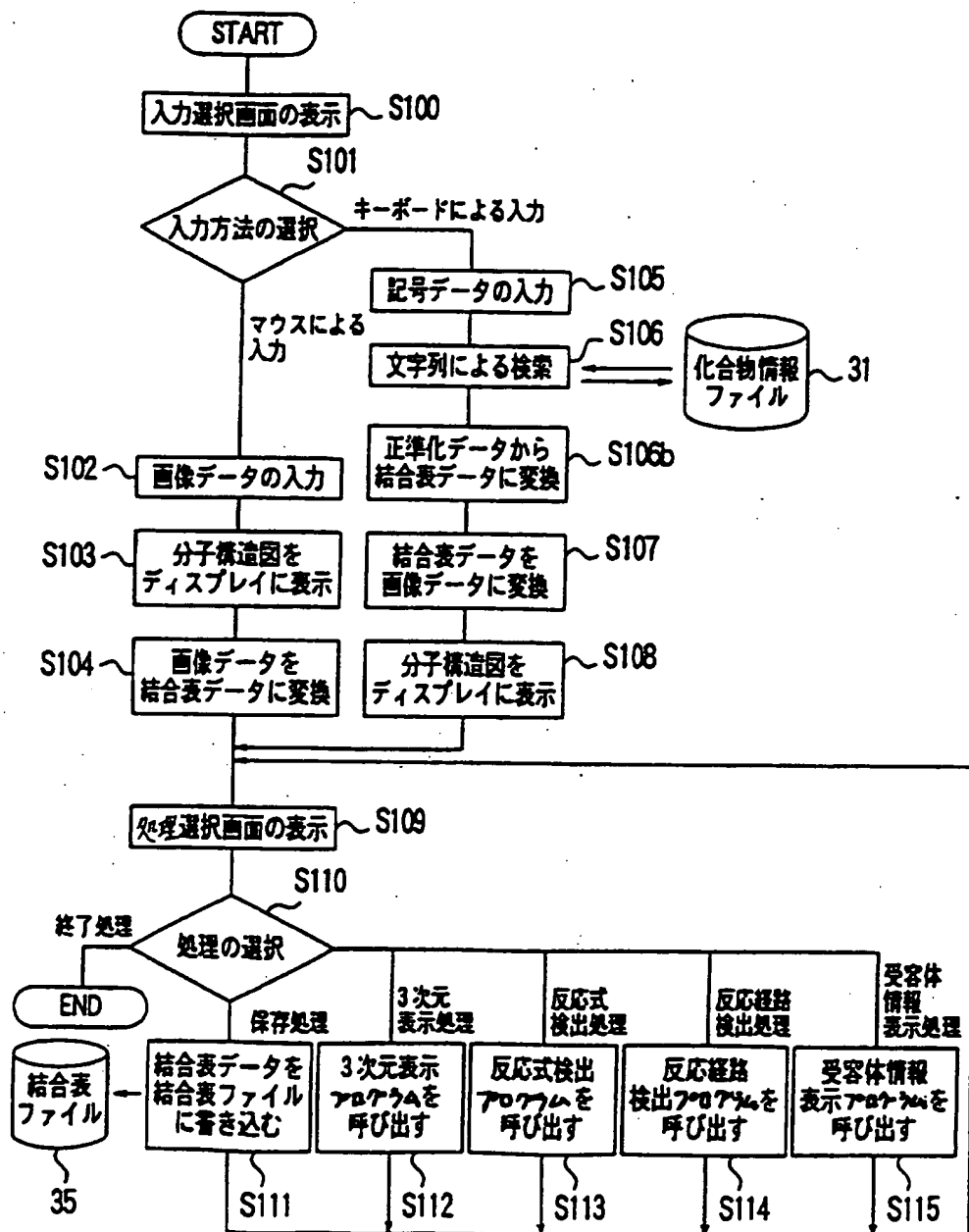


図10A 画像データ 80c (分子構造図)

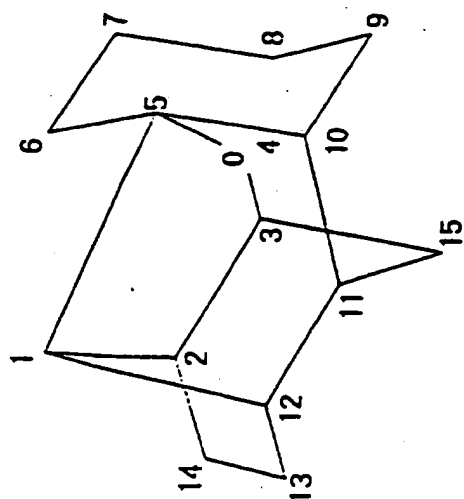


図10B 画像データ 80d (分子構造図)

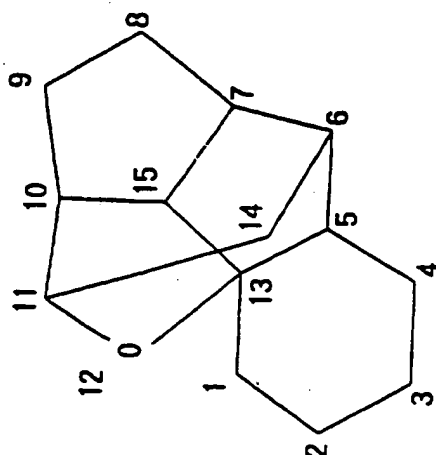


図10C 正準化データ 82c

1-1-1-01-2-2-3-3-4-5-6-7-8-9- /6-8/7-
10/10-14/11-15/12-13/

図 1 2

3次元表示 プログラム 2 4

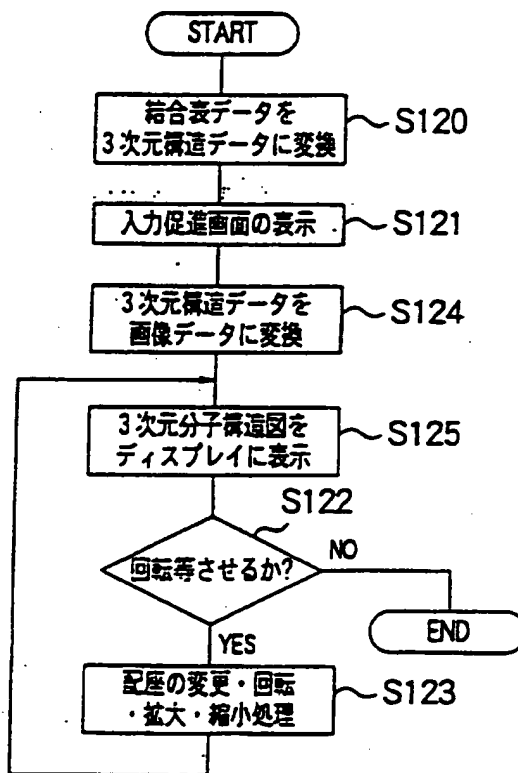


図 1 3

反応式検出プログラム 25

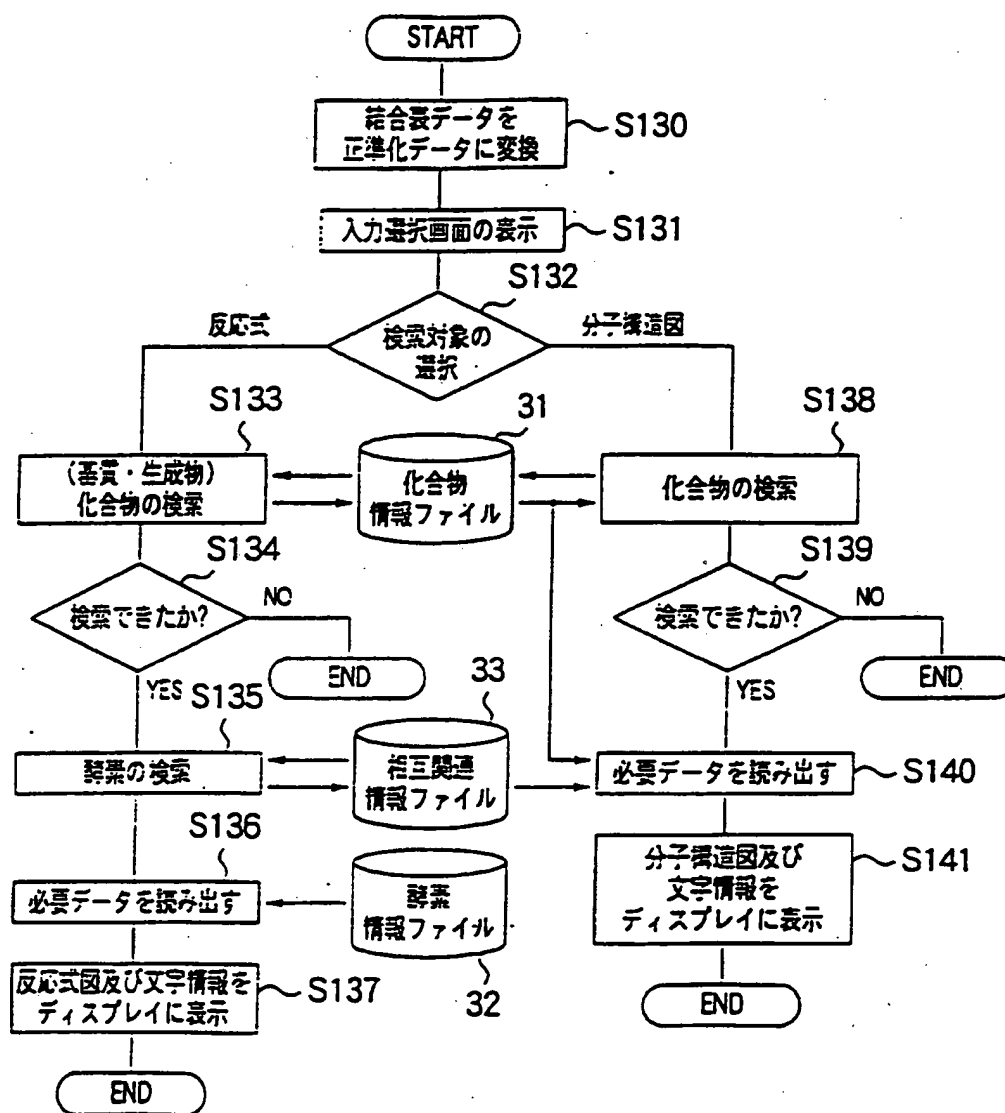


図 14

反応経路検出プログラム 27

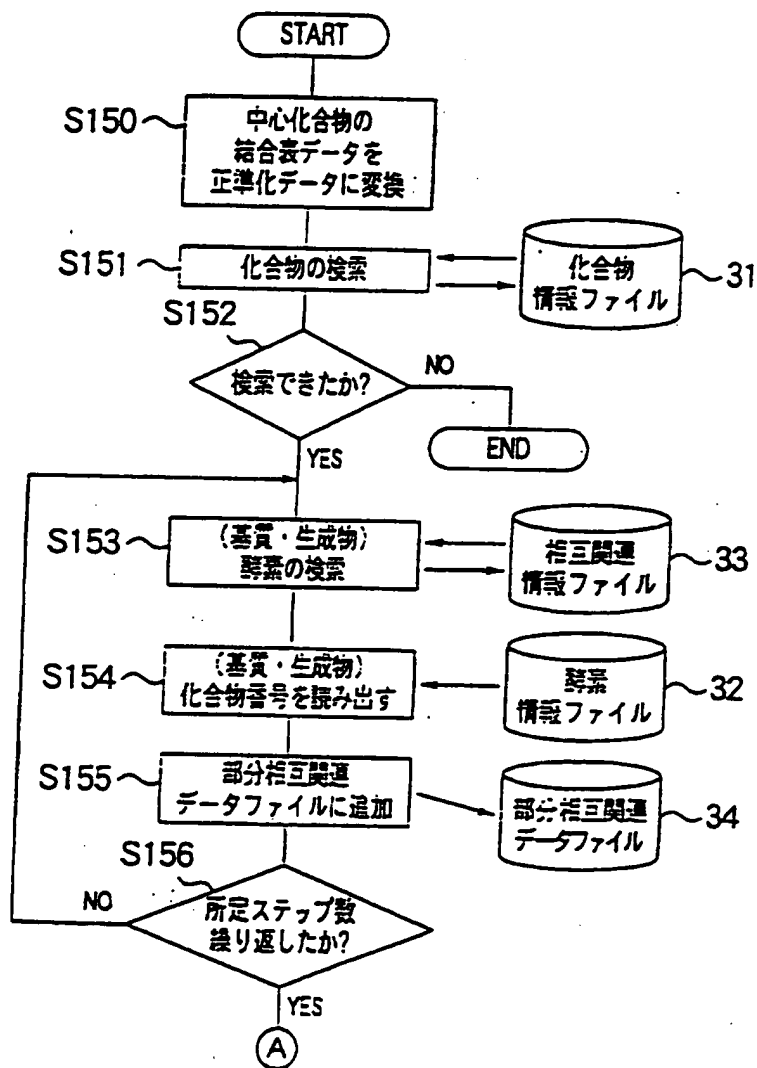


図 15

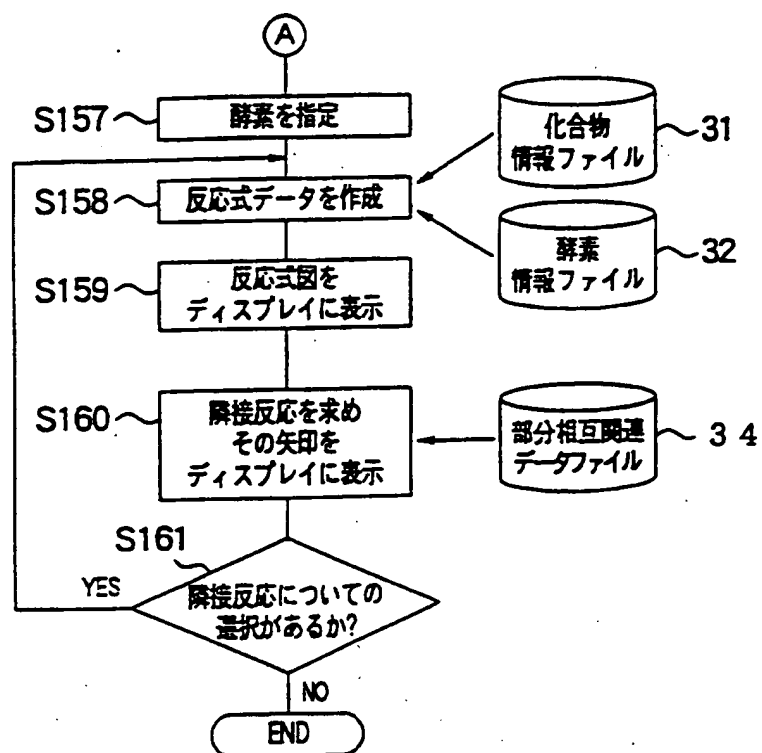


図 16

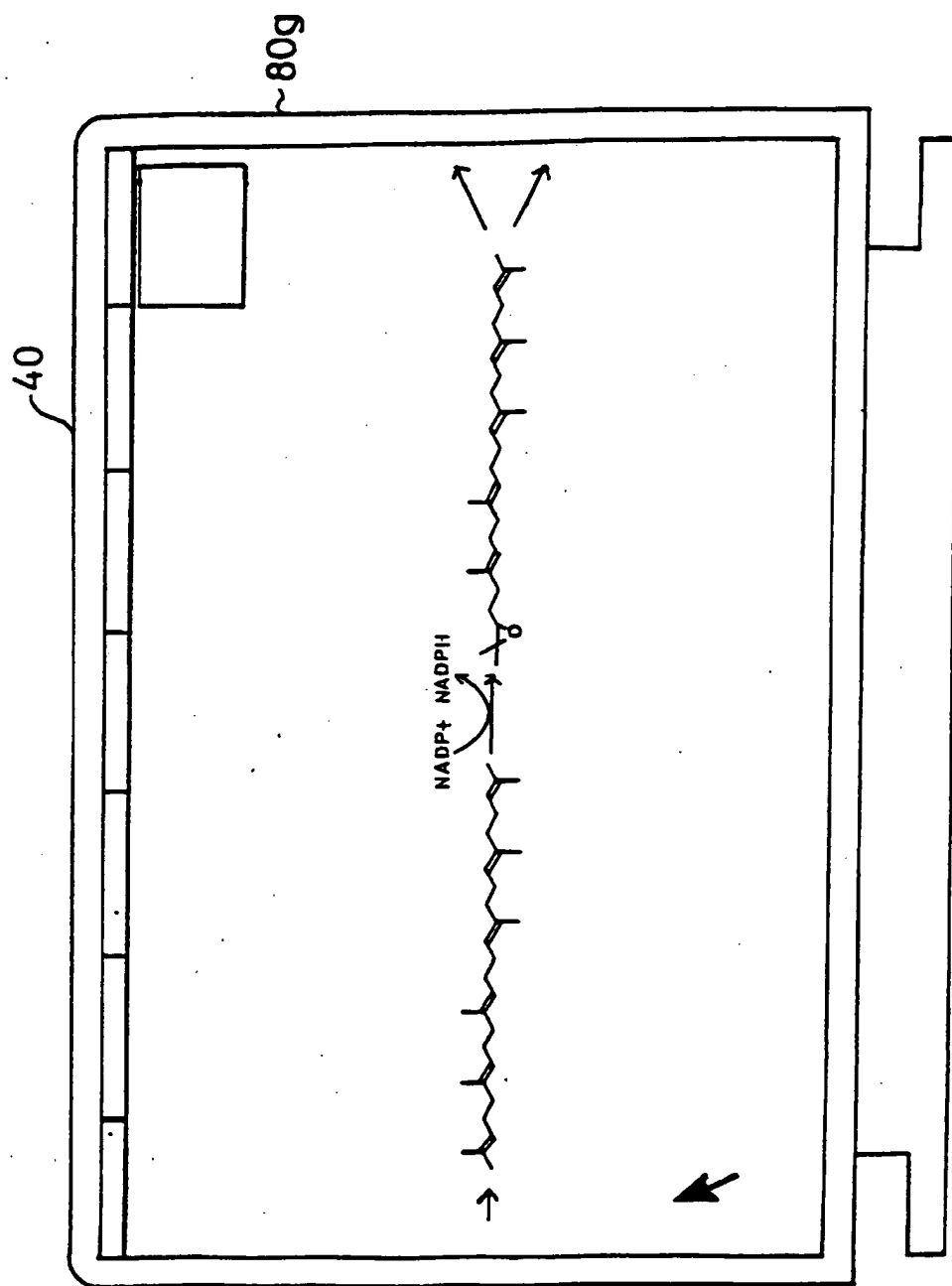
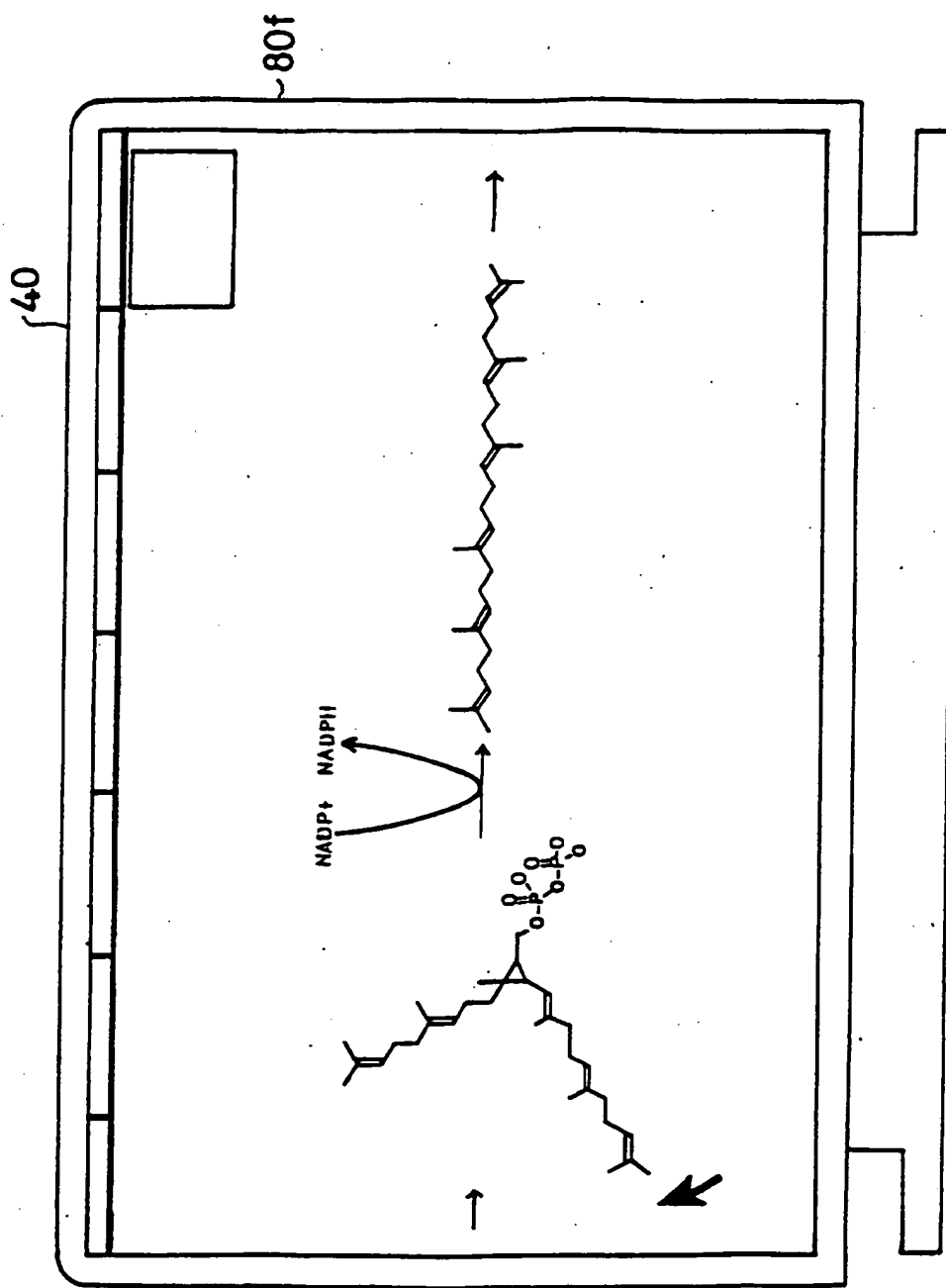


図 17



結合表 81

原子テーブル

81c

図 18 A

入力番号	X座標	Y座標	元素名	属 性	原子数
1	3.0386	-2.4082	C		7
2	3.0386	-0.9082	C		
3	1.7396	-0.1582	C		
4	0.4405	-0.9082	N		7
5	0.4405	-2.4082	C		
6	1.7396	-3.1582	C		
7	4.3376	-3.1582	C		

原子対テーブル

81d

図 18 B

結合原子対		結合種	構 造
1	2	2	環
2	3	1	環
3	4	2	環
4	5	1	環
5	6	2	環
6	1	1	環
1	7	1	鎖

図 19

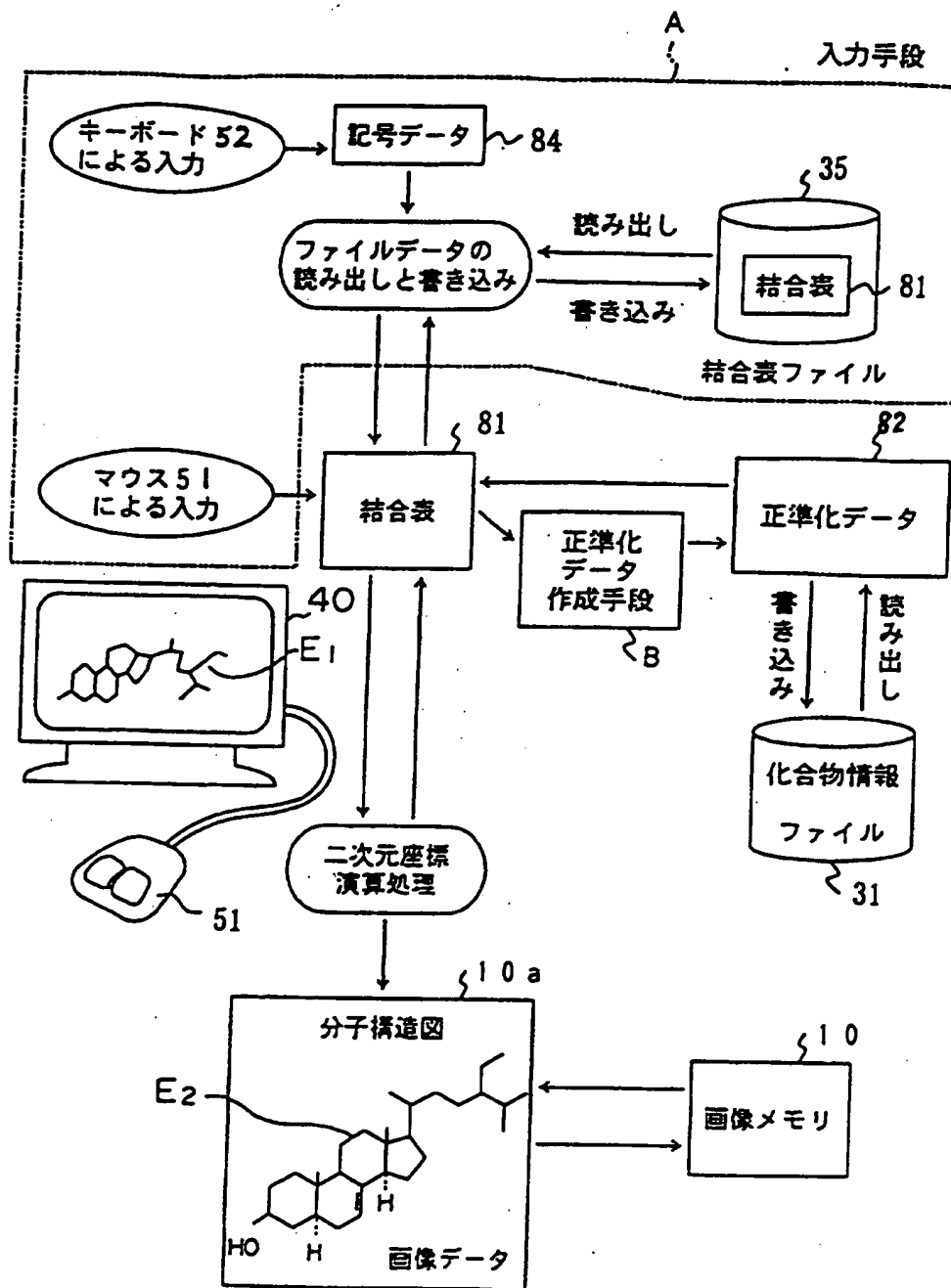


図 20

メインルーチン 91a

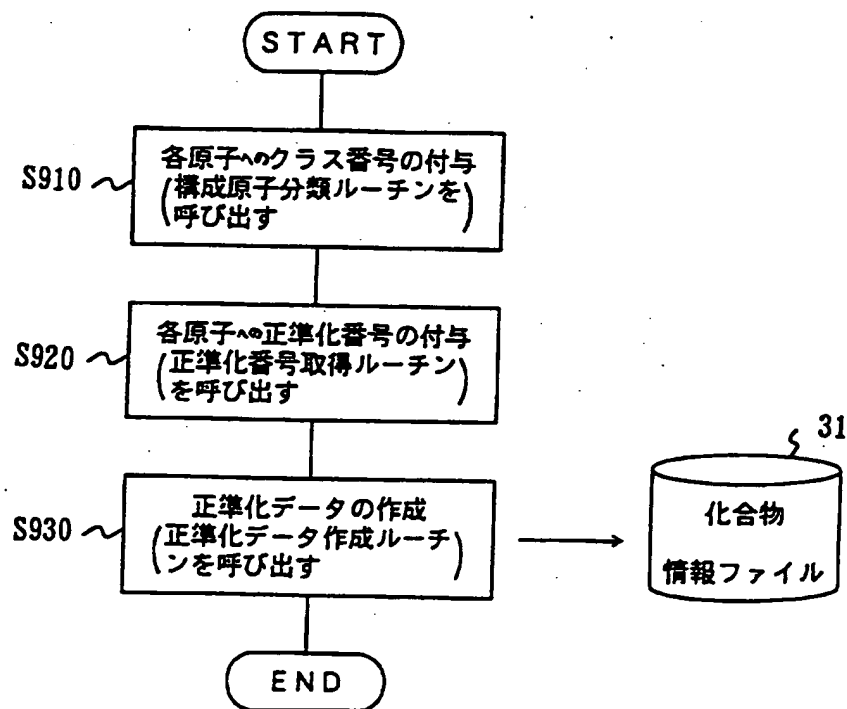
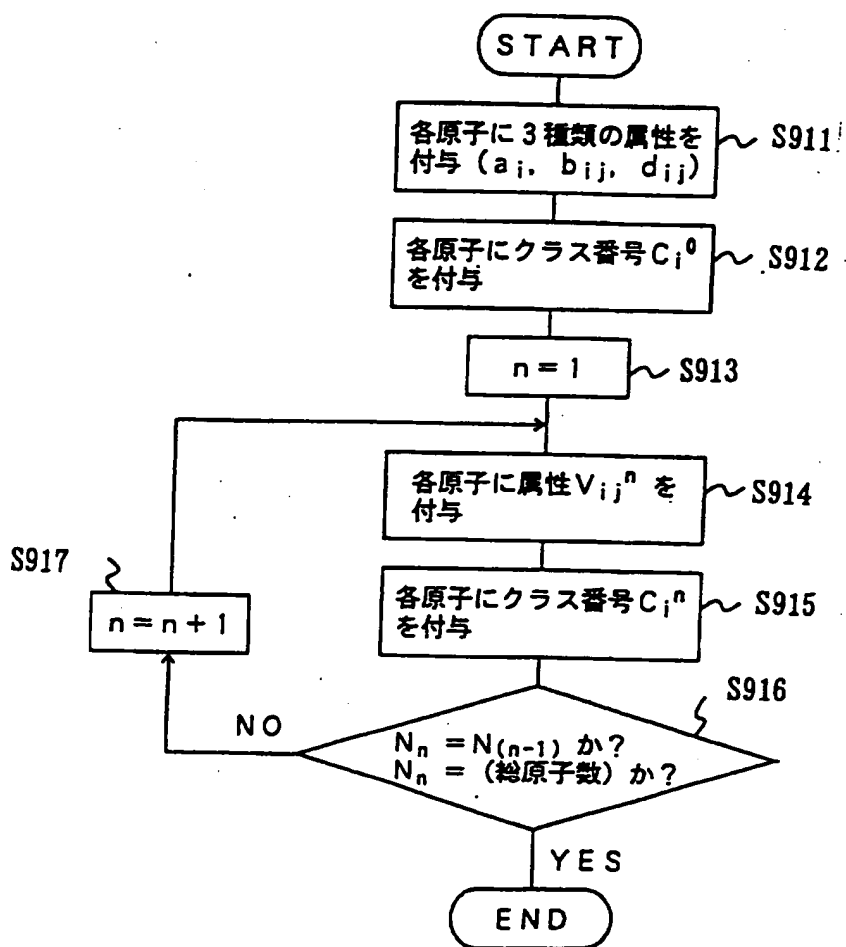


図 2 1

構成原子分類ルーチン 91b



結合表 81

原子テーブル

81e

図 2 2 A

入力番号	X座標	Y座標	元素名	属 性	原子数
1	-	-	C		8
2	-	-	C		
3	-	-	N		結合数
4	-	-	C		8
5	-	-	C		
6	-	-	C		
7	-	-	C		
8	-	-	C		

原子対テーブル

81f

図 2 2 B

結合原子対		結合種	構 造
1	2	1	
2	3	2	
3	4	1	
4	5	1	
5	6	1	
6	1	1	
5	7	1	
1	8	1	

図 23

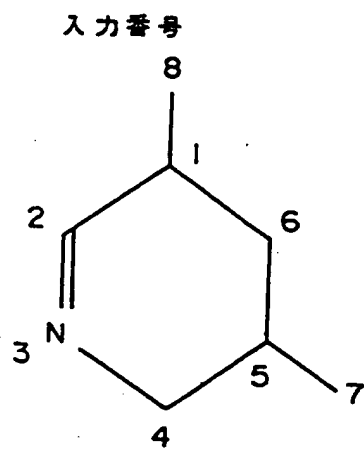


図 24 A

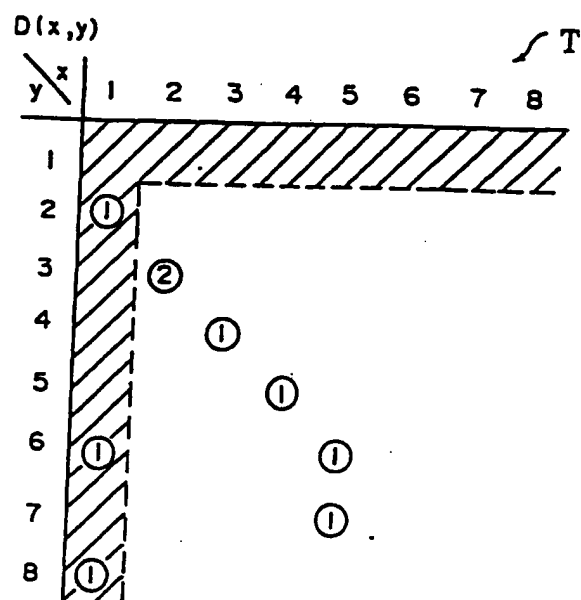
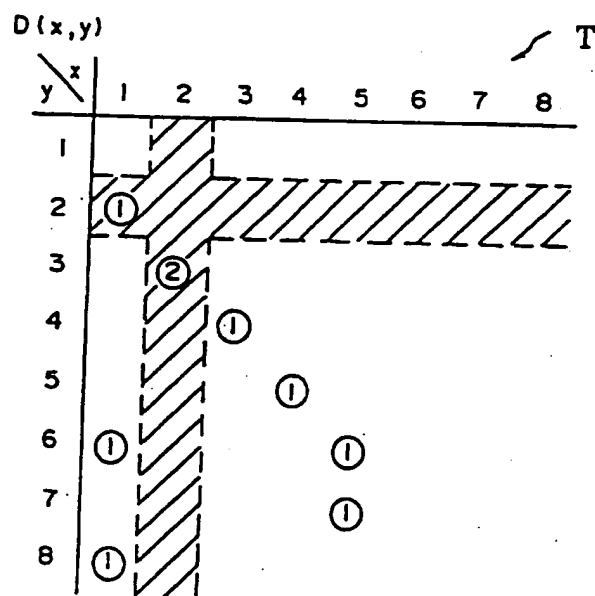


図 24 B



25

$n = 0$

i	a_i	b_{ij}	d_{ij}	C_i^0
1	6	(3, 0, 0, 0)	(3, 2, 3, 0)	5
2	6	(1, 1, 0, 0)	(2, 3, 2, 2)	2
3	7	(1, 1, 0, 0)	(2, 2, 4, 0)	6
4	6	(2, 0, 0, 0)	(2, 3, 2, 2)	3
5	6	(3, 0, 0, 0)	(3, 2, 3, 0)	5
6	6	(2, 0, 0, 0)	(2, 4, 2, 0)	4
7	6	(1, 0, 0, 0)	(1, 2, 2, 3)	1
8	6	(1, 0, 0, 0)	(1, 2, 2, 3)	1

$N_0 = 6$

図 26 A

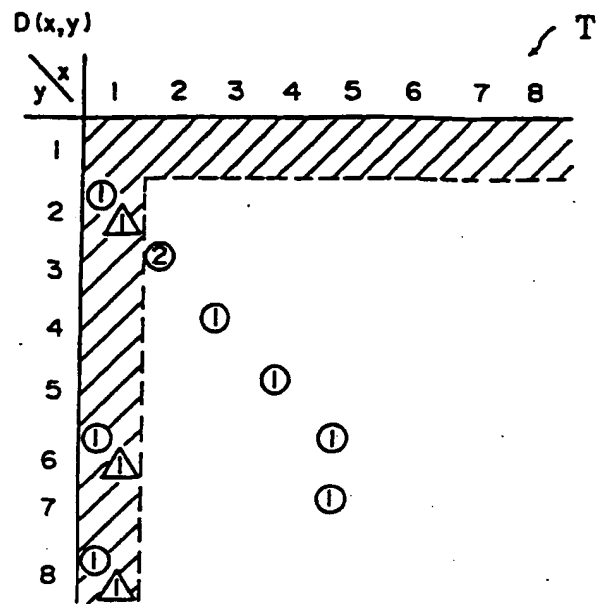
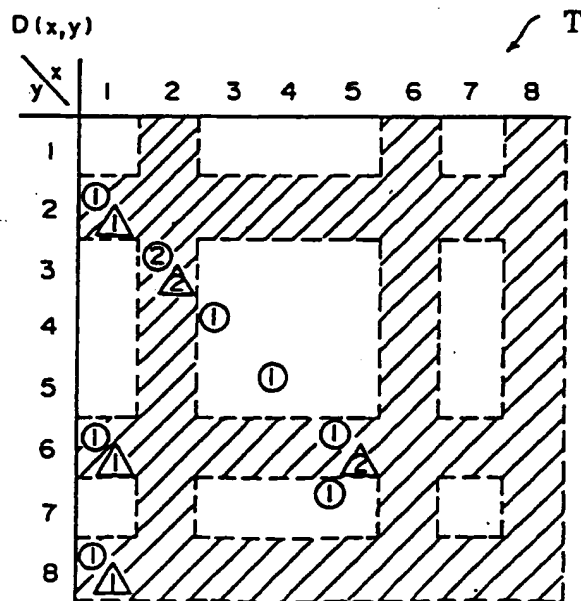
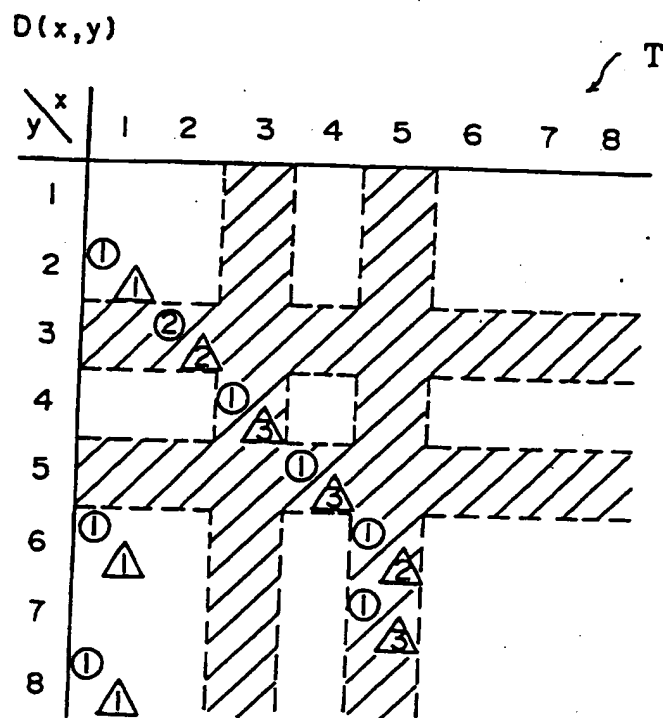


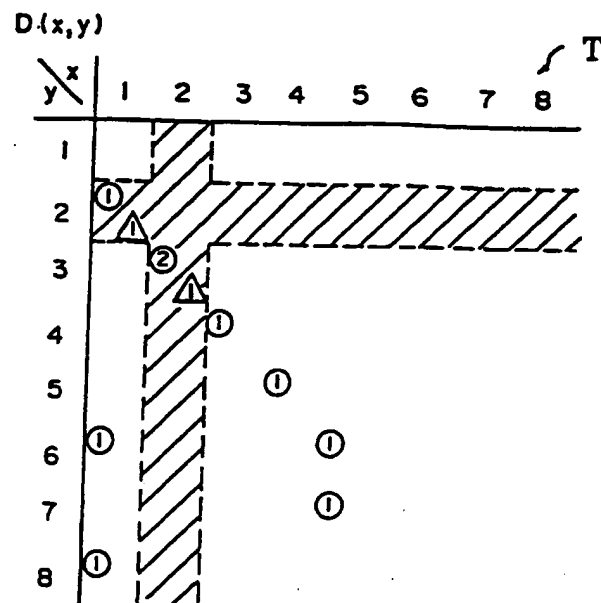
図 26 B



27



28 A



28 B

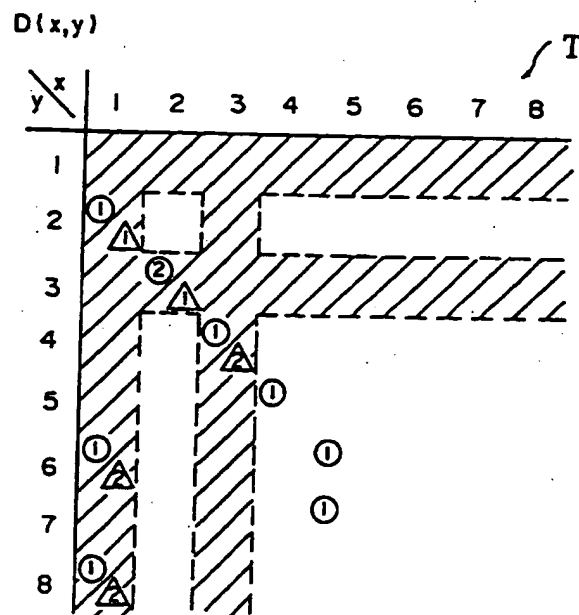


図 29 A

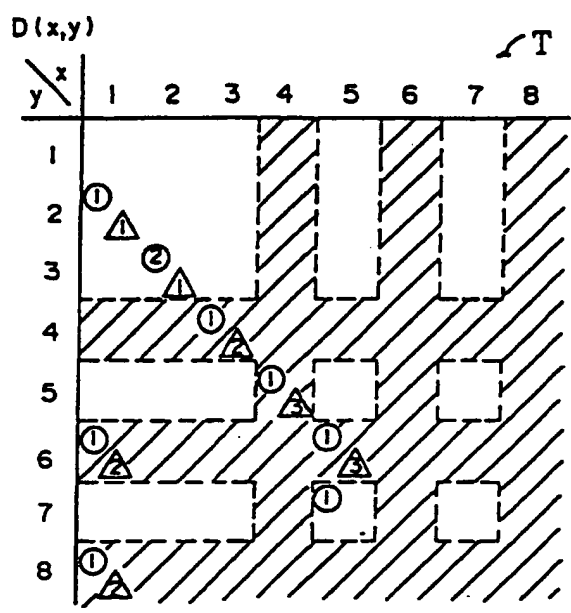


図 29 B

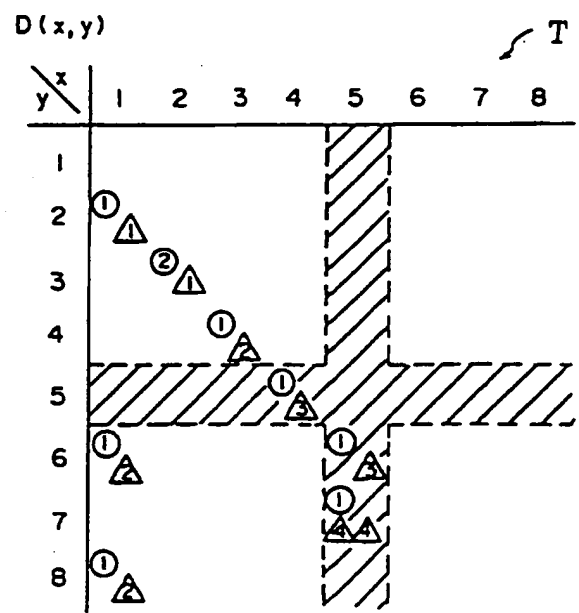


図 30 A

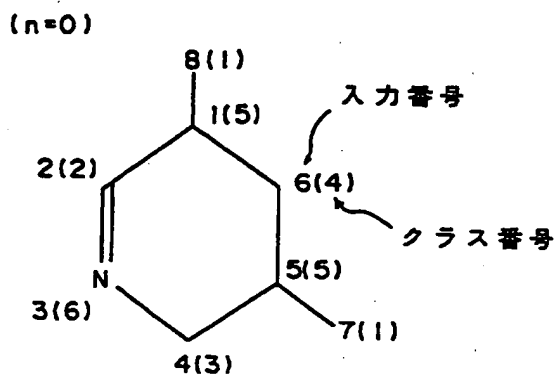


図 30 B

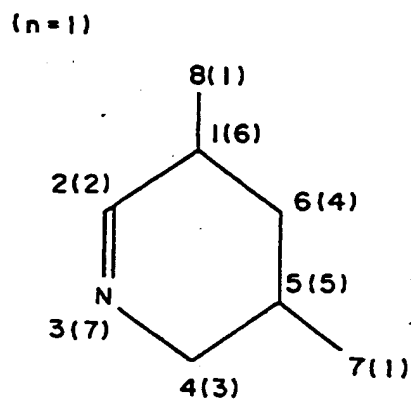


図 30 C

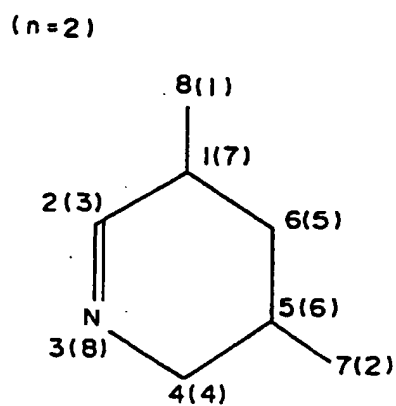


図 3 1

 $n = 1$

i	V_{ij}^1	C_i^1
1	(1, 1, 0, 1, 0, 0)	6
2	(0, 0, 0, 0, 1, 1)	2
3	(0, 1, 1, 0, 0, 0)	7
4	(0, 0, 0, 1, 1, 0)	3
5	(1, 0, 1, 1, 0, 0)	5
6	(0, 0, 0, 0, 2, 0)	4
7	(0, 0, 0, 0, 1, 0)	1
8	(0, 0, 0, 0, 1, 0)	1

 $N_1 = 7$

□ 3 2

 $n = 2$

i	V_{ij}^2	C_i^2
1	(1, 1, 0, 1, 0, 0, 0)	7
2	(0, 0, 0, 0, 0, 1, 1)	3
3	(0, 1, 1, 0, 0, 0, 0)	8
4	(0, 0, 0, 0, 1, 0, 1)	4
5	(1, 0, 1, 1, 0, 0, 0)	6
6	(0, 0, 0, 0, 1, 1, 0)	5
7	(0, 0, 0, 0, 1, 0, 0)	2
8	(0, 0, 0, 0, 0, 1, 0)	1

 $N_2 = 8$

図 3 3

正準化 号付与ルーチン 91c

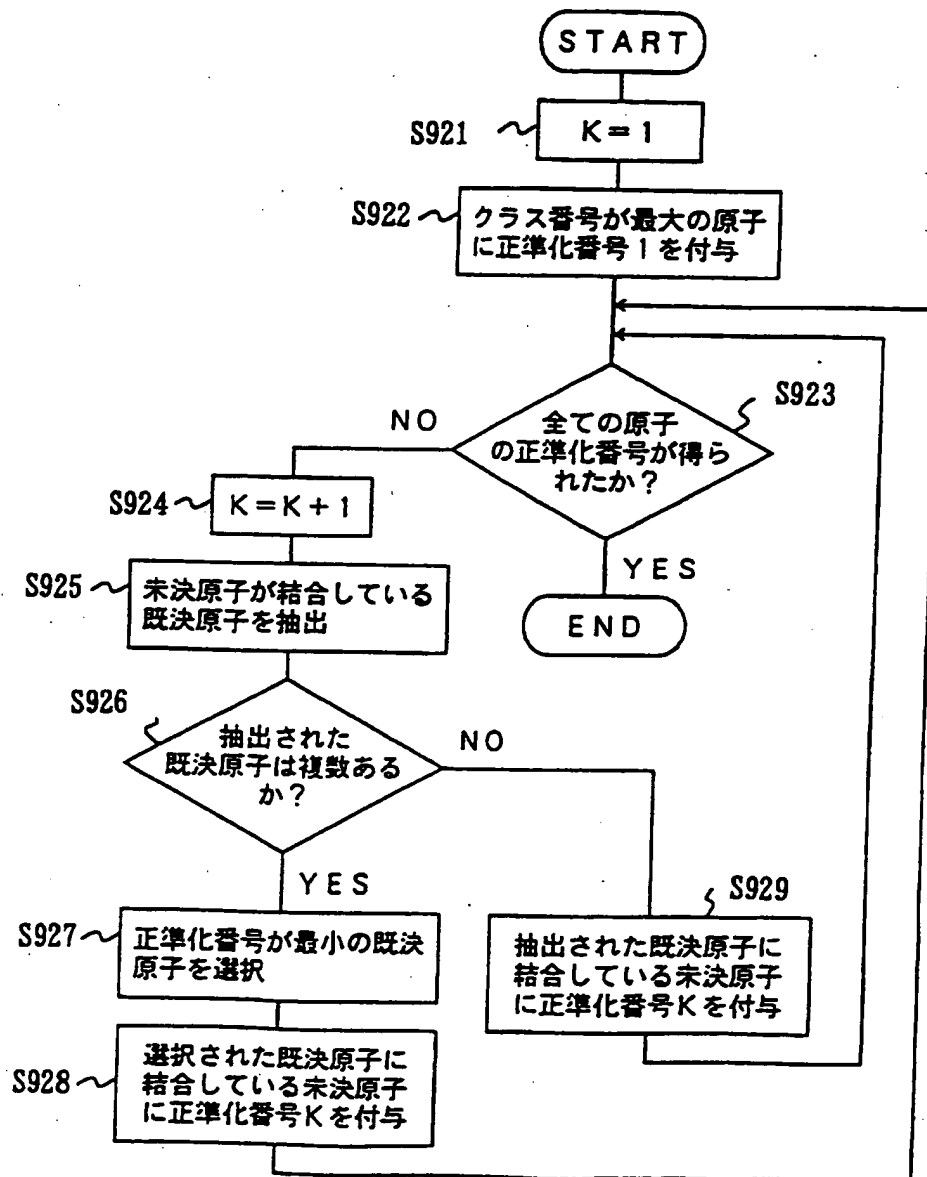
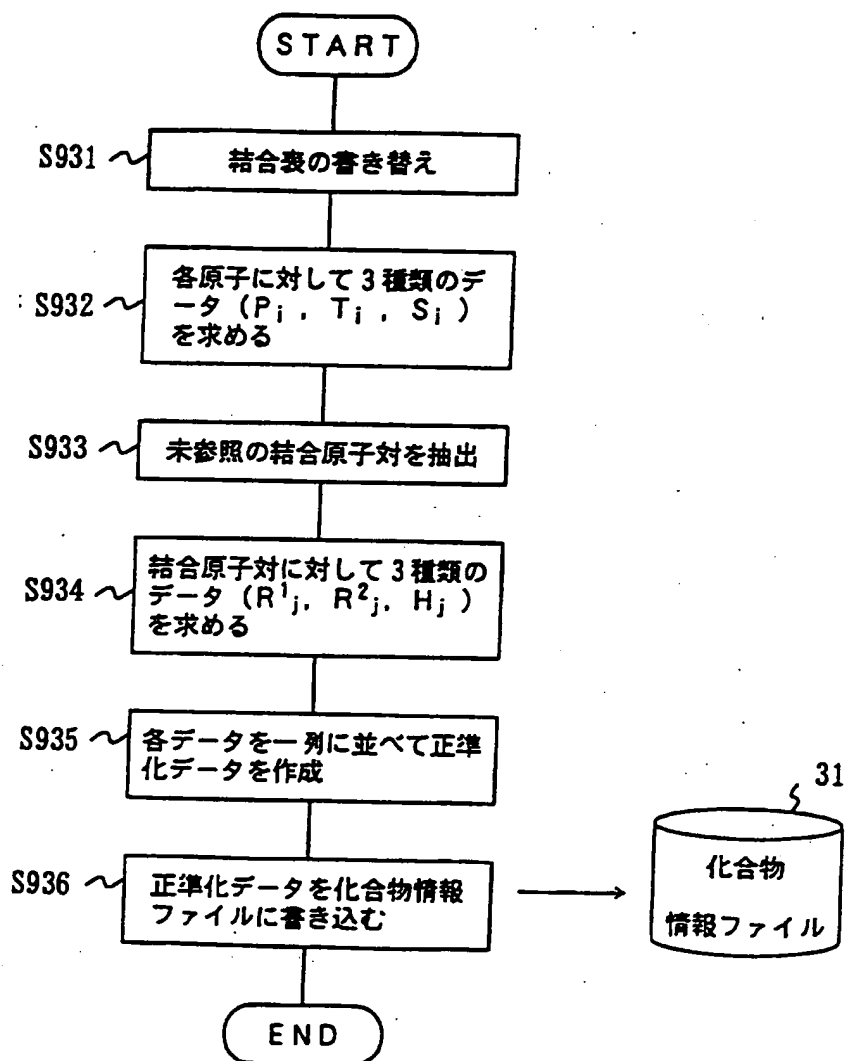


図 3 5

正準化データ作成ルーチン 91d



結合表 81

原子テーブル

81g

図 3 6 A

正準化番号	X座標	Y座標	元素名	属 性	原子数
1	-	-	N		8
2	-	-	C		結合数
3	-	-	C		
4	-	-	C		8
5	-	-	C		
6	-	-	C		
7	-	-	C		
8	-	-	C		

原子対テーブル

81h

図 3 6 B

結合原子対		結合種	構 造
5	3	1	
3	1	2	
1	2	1	
2	4	1	
4	6	1	
6	5	1	
4	7	1	
5	8	1	

図 37

K		P_K	T_K	S_K
1		N
2		1	—	C
3		1	=	C
4		2	—	C
5		3	—	C
6		4	—	C
7		4	—	C
8		5	—	C
ℓ	R_{ℓ}^1	R_{ℓ}^2	H_{ℓ}	
1	5	6	—	

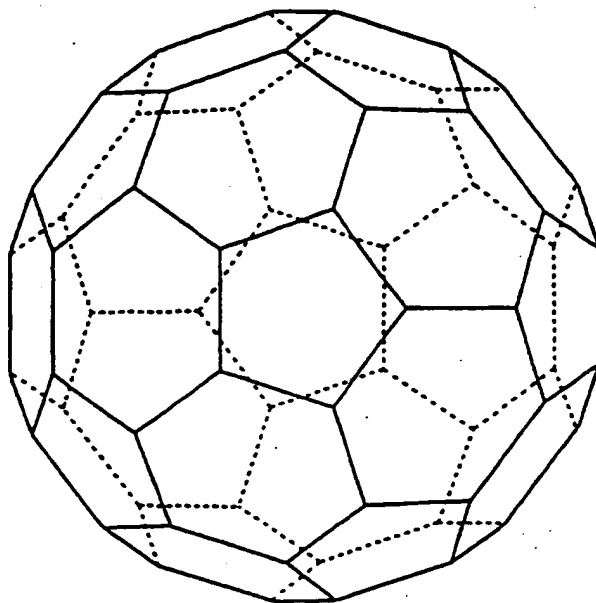
図 3 8 A 画像データ 80e (C₆₀の分子構造図)

図 3 8 B 正準化データ 82d

1%1%1%2%2%3%3%4%4%5%5%6%6%7%8%9%10%1
 1%12%13%14%15%15%16%17%17%18%19%19%2
 1%21%23%24%25%26%27%28%29%30%31%32%3
 3%33%35%36%36%38%39%41%43%44%45%46%4
 7%48%49%50%53%57%/7%12/8%10/9%14/11%
 13/16%24/18%27/20%23/20%30/22%26/22%
 32/25%28/29%31/34%35/34%43/37%38/37%
 46/39%40/40%44/41%42/42%47/45%48/49%
 50/51%52/51%53/52%57/54%55/54%56/55%
 58/56%59/58%60/59%60/

図 3 9

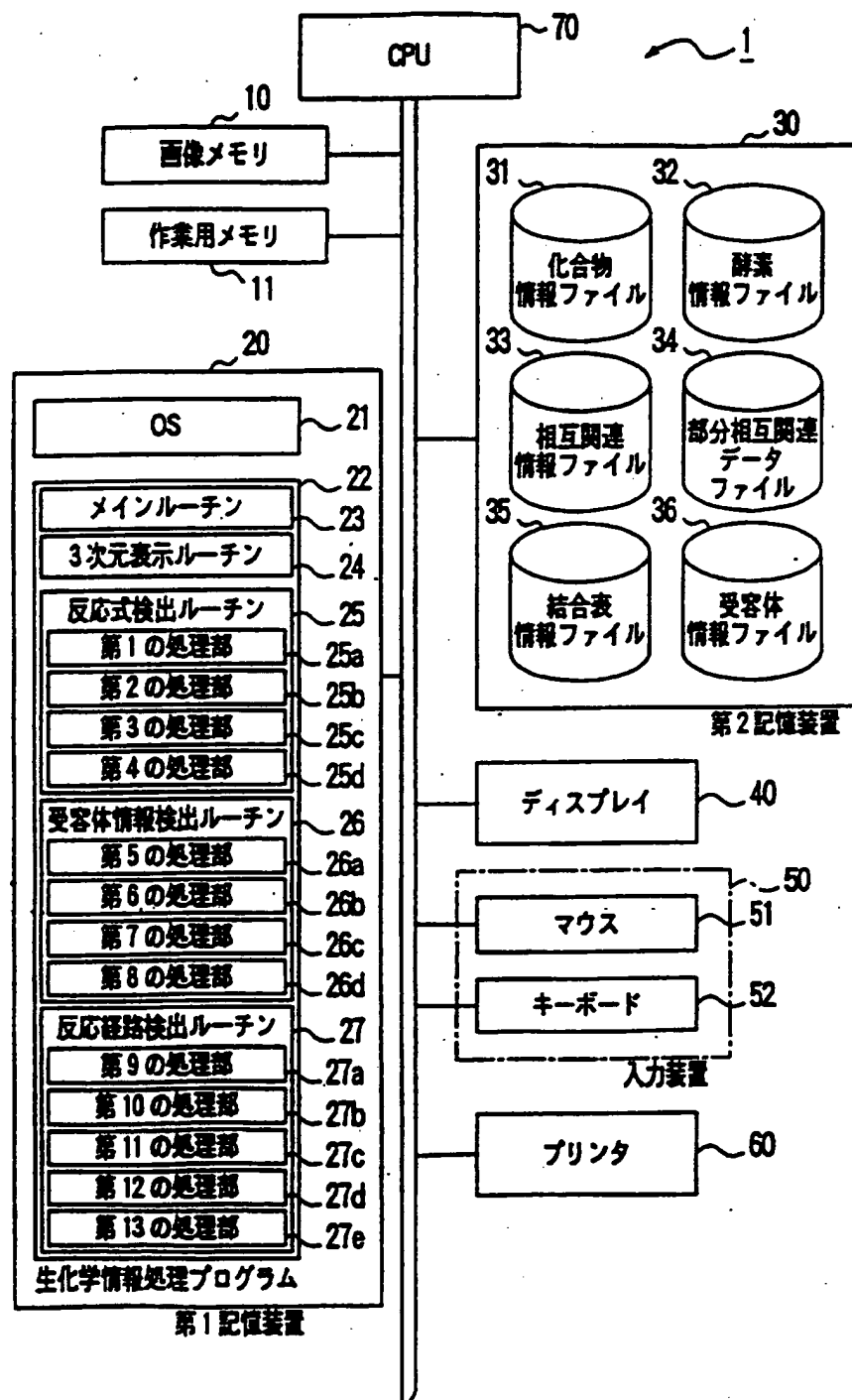


図 40

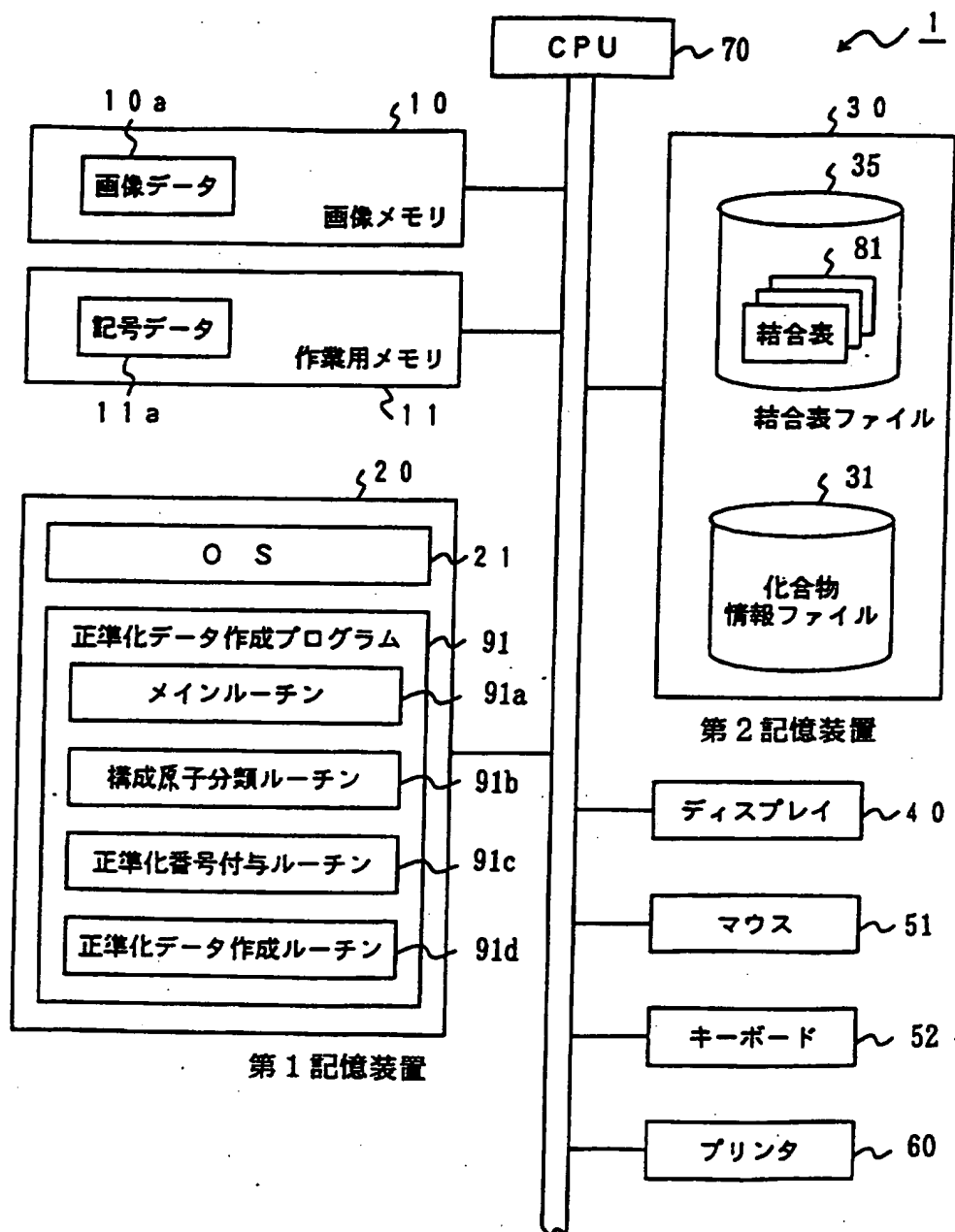


図 4 1

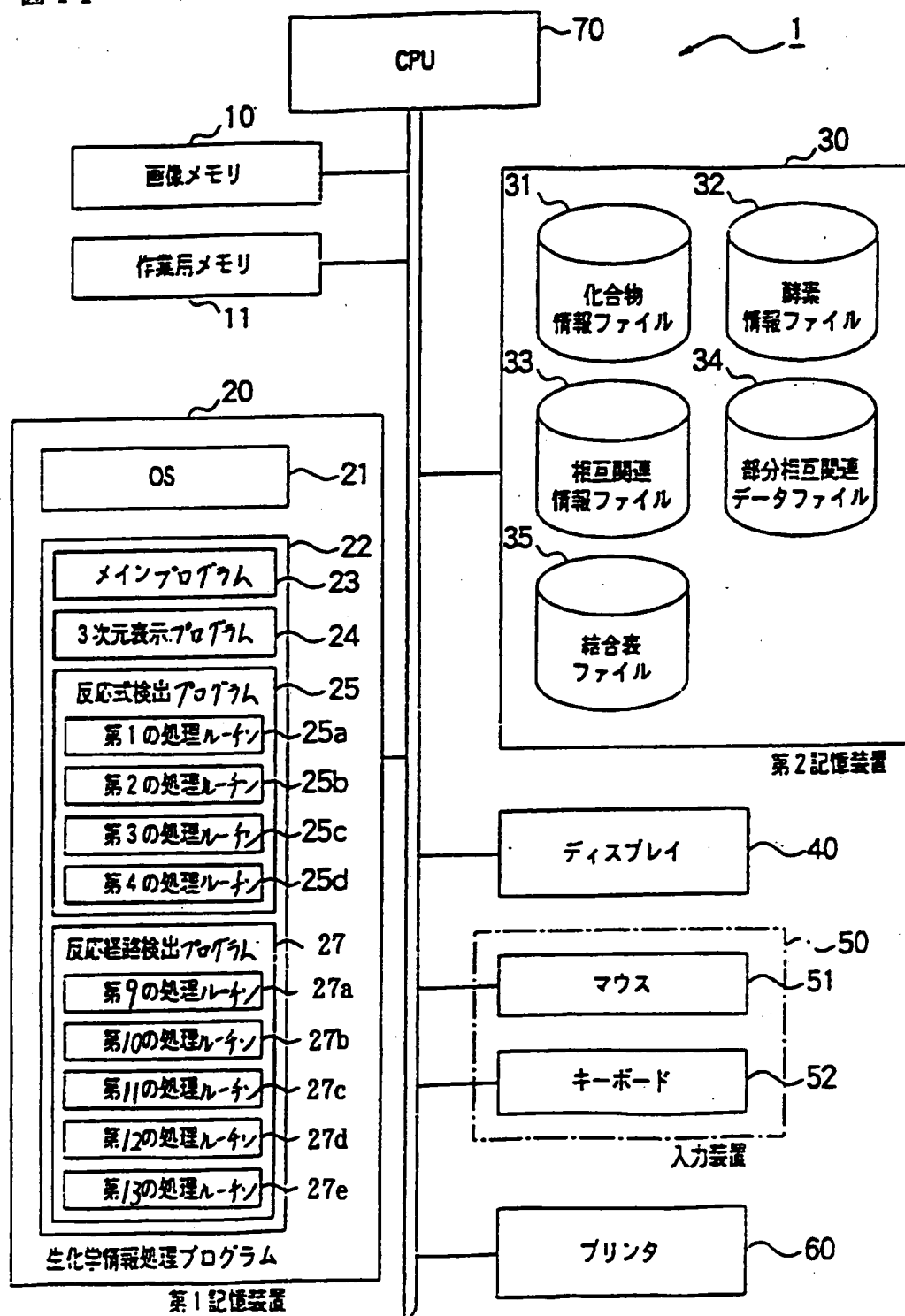


図 4 2

相互関連情報ファイル

33

化合物番号	(基質) 酵素番号	(生成物) 酵素番号	(阻害) 酵素番号
C ₁	E ₁		
C ₂	E ₂	E ₁	
C ₃	E ₃ , E ₄	E ₂	
C ₄		E ₃	
C ₅	E ₅		
C ₆	E ₆	E ₅	E ₄
C ₇		E ₆	

図 4 3

メインプログラム 23

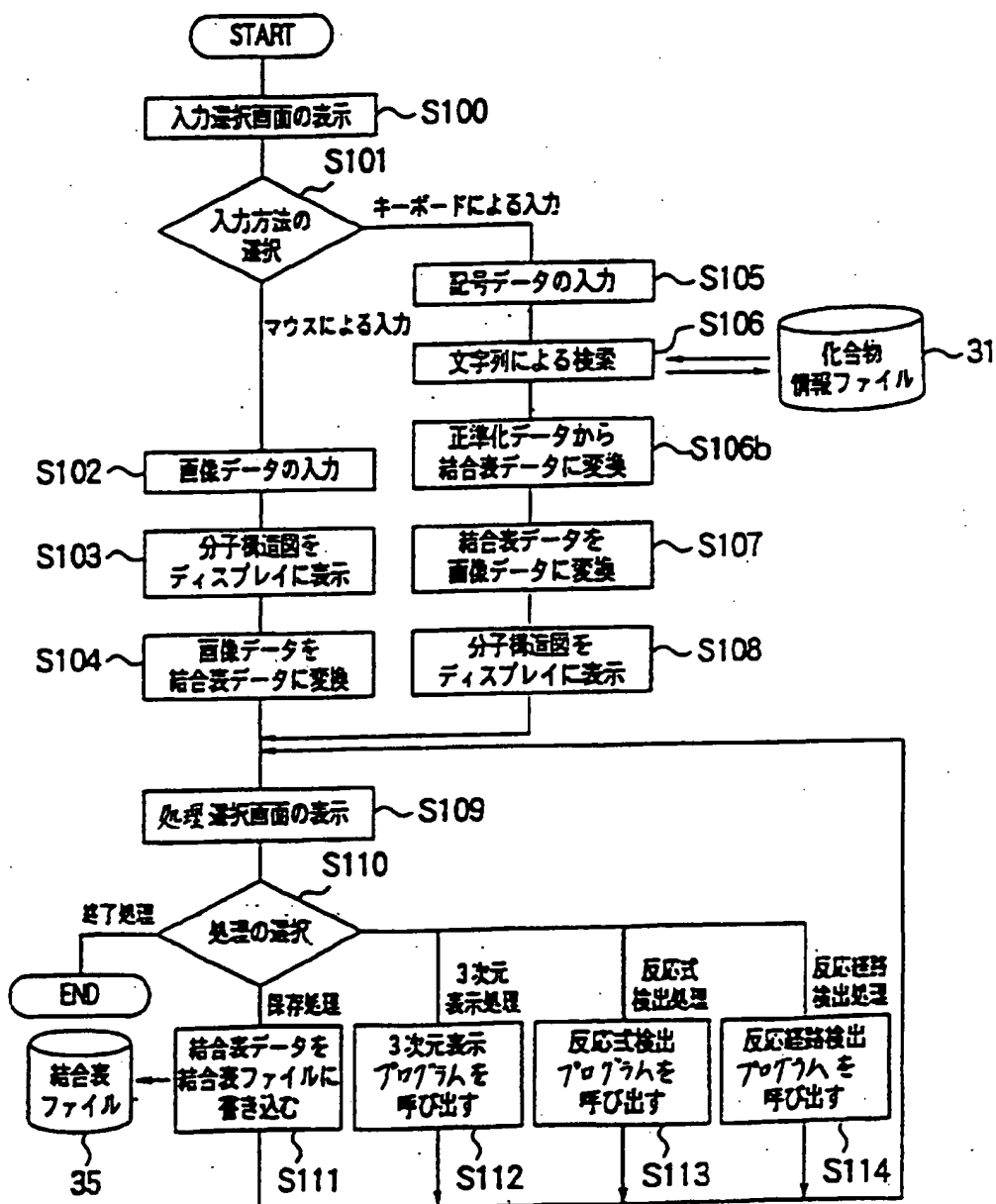


図 4 4

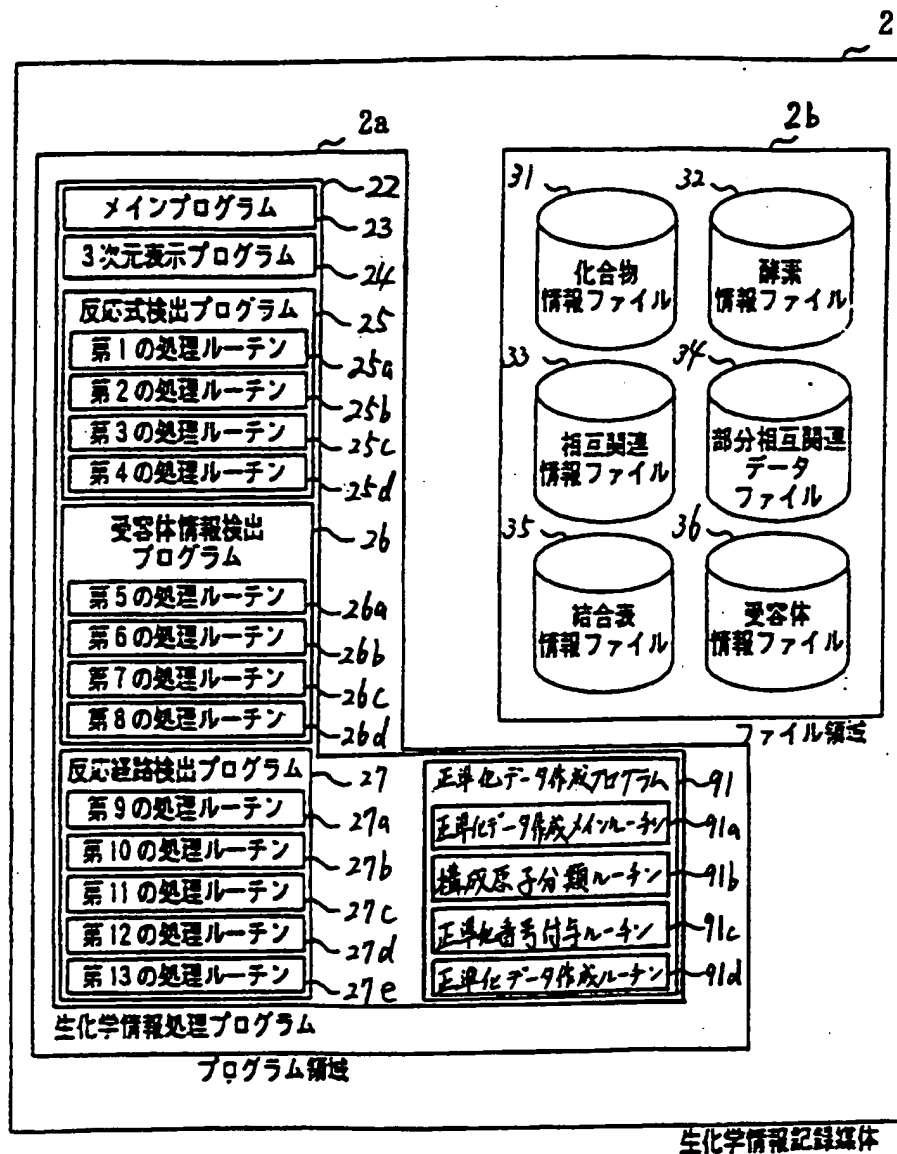


図 4 5

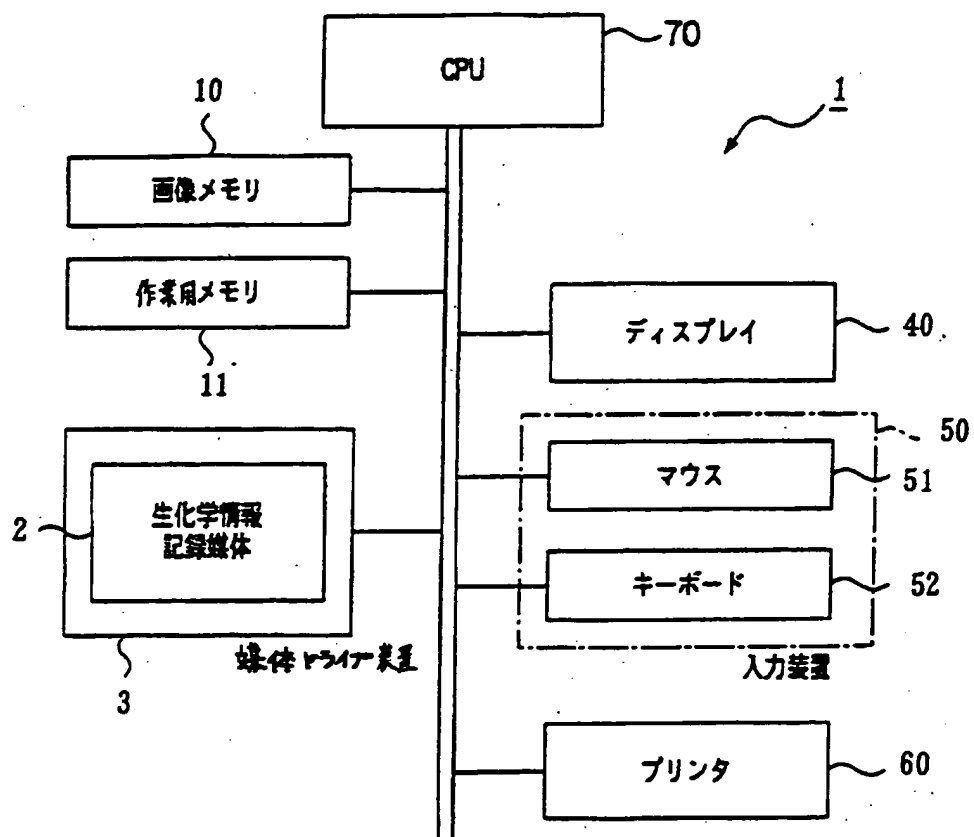


図 4 6

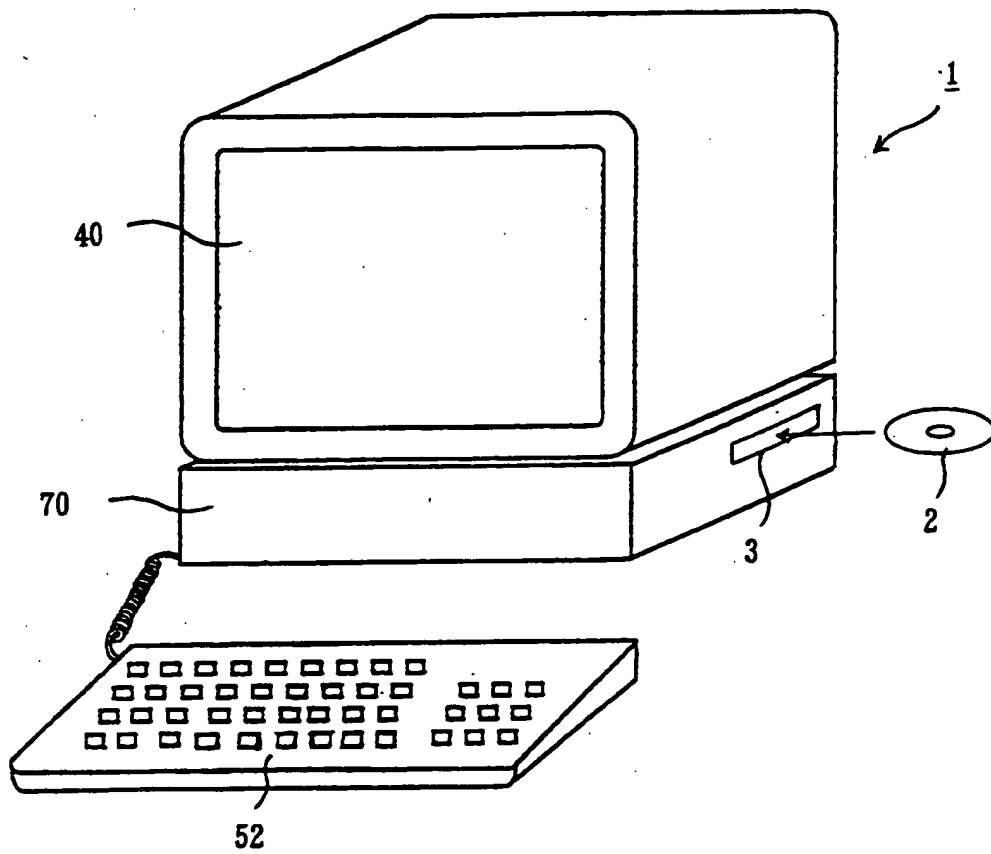


図 4 7

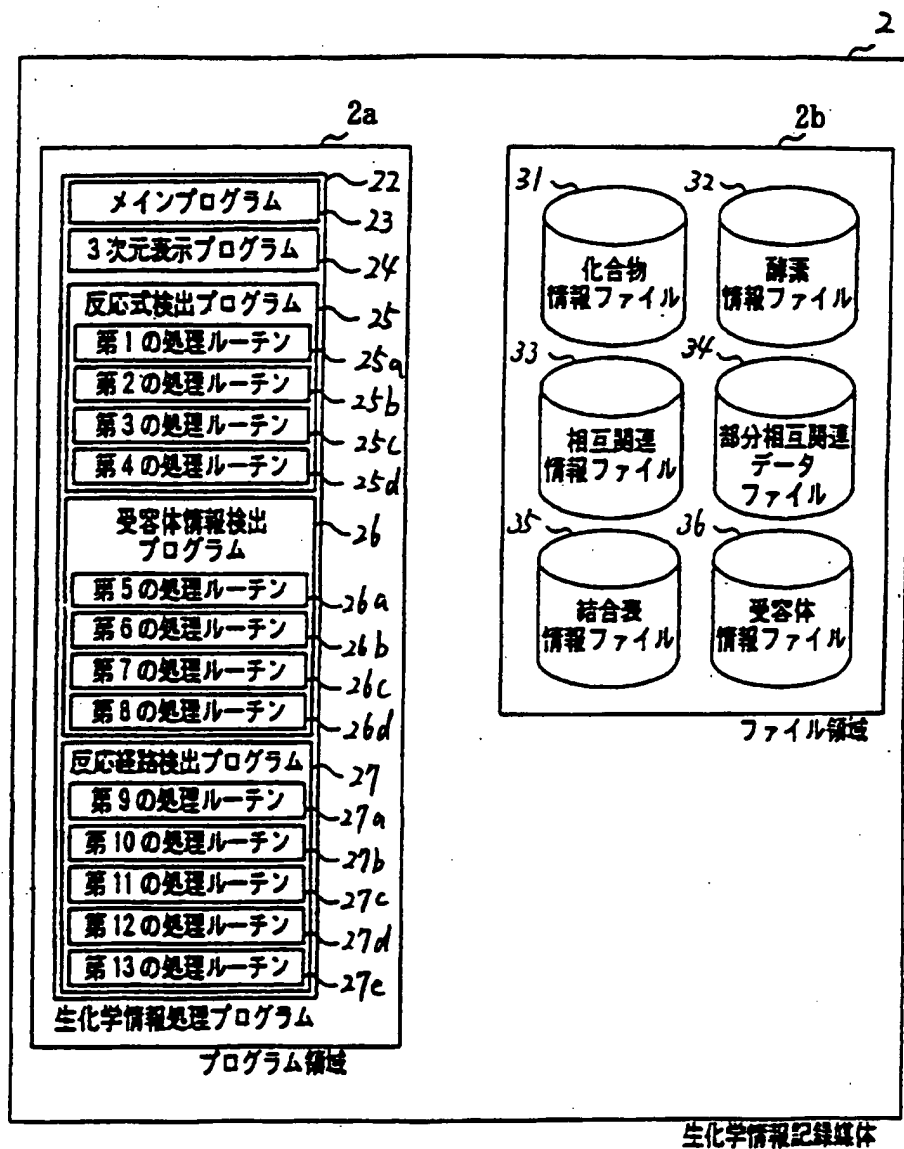
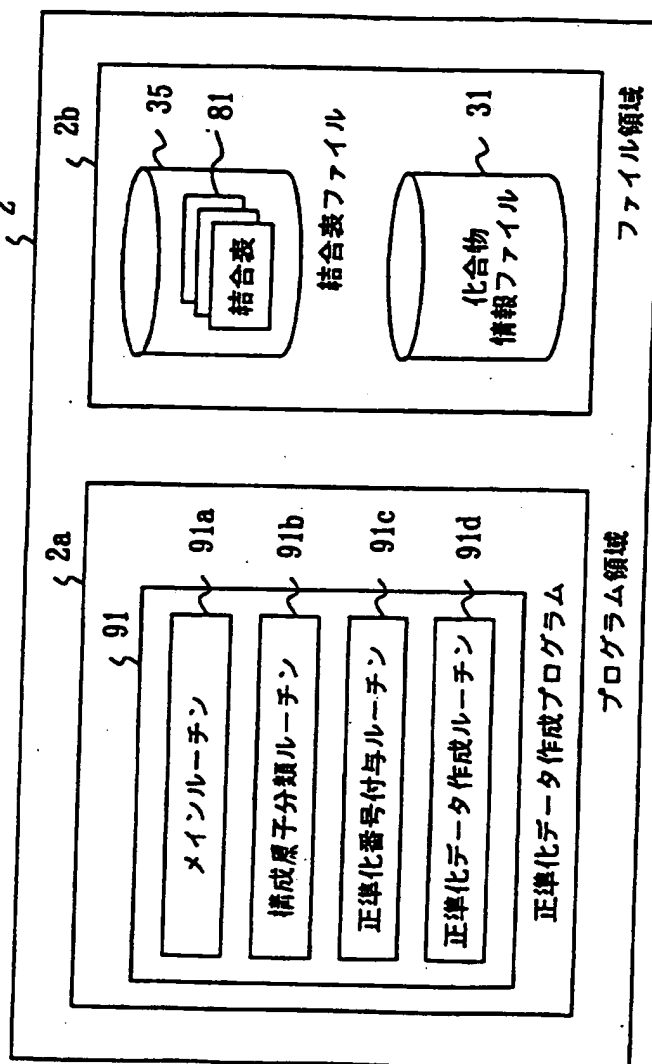


図 48



標準化データ作成用記録媒体

図 4 9

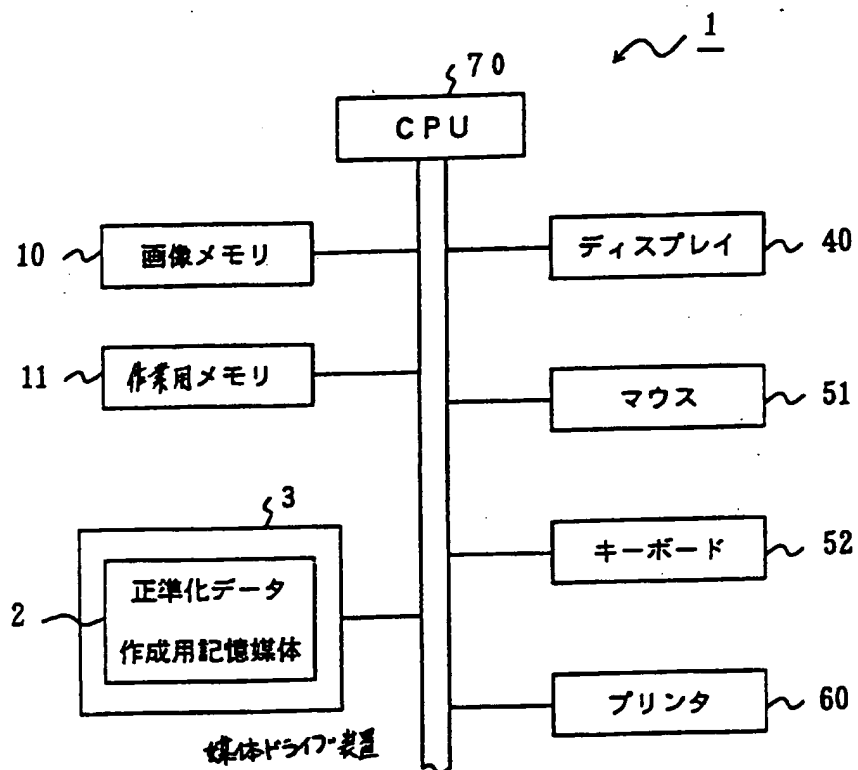
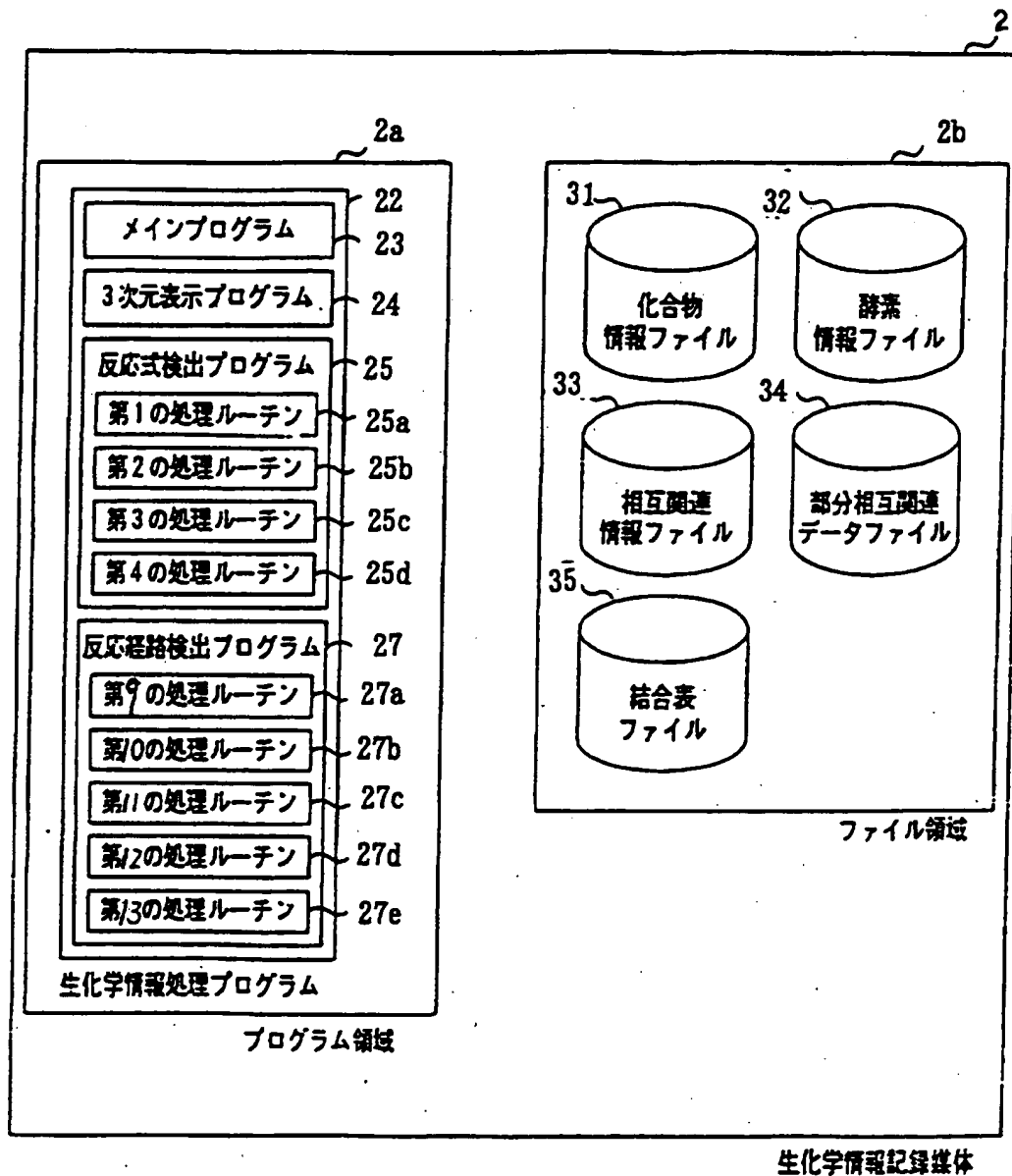


図 50



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP96/00709

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl⁶ G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int. Cl⁶ G06F17/30

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho

1926 - 1996

Kokai Jitsuyo Shinan Koho

1971 - 1996

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JICST (Biochemistry, Data base)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KARP, P.D. "Representing, Analyzing, and Synthesizing Biochemical Pathways." IEEE Expert, 1994, Vol. 9, No. 2, pages 11 to 21	1 - 14
A	BARCZA, S., KELLY, L.A., LENZ, C.D. "Computerized retrieval of information on biosynthesis and metabolic pathways." J. Chem. Inf. Comput. Sci., 1990, Vol. 30, No. 3, pages 243-251	1 - 14

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"T" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"A" document member of the same patent family

Date of the actual completion of the international search

June 11, 1996 (11. 06. 96)

Date of mailing of the international search report

June 25, 1996 (25. 06. 96)

Name and mailing address of the ISA/

Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP96/00709

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 15 - 22
because they relate to subject matter not required to be searched by this Authority, namely:
Inventions as set forth in claims 15 to 22 are considered to be computer programs and computer program is a subject matter for which the international search is not required under the provisions of Rule 42(6) of the Regulations under the Law concerning international applications, etc. pursuant to the Patent Cooperation Treaty.
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl.⁸. G06F17/30

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl.⁸. G06F17/30

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1926-1996年

日本国公開実用新案公報 1971-1996年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICST (生化学、データベース)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	KARP, P. D. "Representing, Analyzing, and Synthesizing Biochemical Pathways." IEEE Expert, 1994, Vol. 9, No. 2, pages 11 to 21	1-14
A	BARCZA, S. KELLY, L. A. LENZ, C. D. "Computerized retrieval of information on biosynthesis and metabolic pathways." J. Chem. Inf. Comput. Sci., 1990, Vol. 30, No. 3, pages 243-251.	1-14

☐ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの

「E」 先行文献ではあるが、国際出願日以後に公表されたもの

「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」 口頭による開示、使用、展示等に言及する文献

「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」 同一パテントファミリー文献

国際調査を完了した日

11. 06. 96

国際調査報告の発送日

25.06.96

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便 号100

東京都千代田区霞が関三丁目4 3号

特許庁審査官 (権限のある職員)

5 L

9 2 8 9

印

野崎 大造

電話番号 03-3581-1101 内線 3564

第I欄 請求の範囲の一部の調査ができないときの意見 (第1ページの1の続き)

法第8条第3項 (PCT 17条(2)(a)) の規定により、この国際調査報告は次の理由により請求の範囲の一部について作成しなかった。

1. ☒ 請求の範囲 15-22 は、この国際調査機関が調査をすることを要しない対象に係るものである。
つまり、
請求の範囲15-22に係る発明は、コンピュータプログラムであると認められ、コンピュータプログラムは、特許協力条約に基づく国際出願等に関する法律施行規則第42条(6)に定められた国際調査を要しない対象であるためである。
2. ☐ 請求の範囲 _____ は、有意義な国際調査をすることができる程度まで所定の要件を満たしていない国際出願の部分に係るものである。つまり、
3. ☐ 請求の範囲 _____ は、従属請求の範囲であってPCT規則6.4(a)の第2文及び第3文の規定に従って記載されていない。

第II欄 発明の単一性が欠如しているときの意見 (第1ページの2の続き)

次に述べるようにこの国際出願に二以上の発明があるとこの国際調査機関は認めた。

1. ☐ 出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求の範囲について作成した。
2. ☐ 追加調査手数料を要求するまでもなく、すべての調査可能な請求の範囲について調査することができたので、追加調査手数料の納付を求めなかった。
3. ☐ 出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったので、この国際調査報告は、手数料の納付のあった次の請求の範囲のみについて作成した。
4. ☐ 出願人が必要な追加調査手数料を期間内に納付しなかったので、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求の範囲について作成した。

追加調査手数料の異議の申立てに関する注意

- ☐ 追加調査手数料の納付と共に出願人から異議申立てがあった。
☐ 追加調査手数料の納付と共に出願人から異議申立てがなかった。